1976

Aspects of quadratic programming with statistical applications

Rebecca Jane Klemm
Iowa State University

77-1457

KLEMM, Rebecca Jane, 1950-
ASPECTS OF QUADRATIC PROGRAMMING WITH STATISTICAL APPLICATIONS.

Iowa State University, Ph.D., 1976
Statistics

Aspects of quadratic programming with

statistical applications

by

Rebecca Jane Klemm

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1976

TABLE OF CONTENTS

CHAPTER I. INTRODUCTION

Consider the nonlinear programming problem

$$\begin{aligned} &\text{minimize} \quad F(y) \\ &\text{subject to} \quad \underline{f}(y) \leq 0 \\ &\qquad\qquad y \in L \end{aligned} \tag{1.1}$$

where $F(y)$ and $\underline{f}(y)$ are arbitrary functions in E^k and L is an arbitrary set in E^k such that $F: E^k \rightarrow E^1$ and $\underline{f}: E^k \rightarrow E^m$. A sufficient and often necessary condition for y^o to be a solution of (1.1) is that there exists a vector $x^o \geq 0$ such that

$$\phi(y^o, x) \leq \phi(y^o, x^o) \leq \phi(y, x^o) \quad \forall x \geq 0. \quad \forall y \in L \tag{1.2}$$

where

$$\phi(y, x) = F(y) + x'\underline{f}(y) .$$

$\phi(y, x)$ is commonly denoted as the Lagrangian function associated with (1.1), and (1.2) is known as the saddle value problem.

In establishing the general equivalence relationship between (1.1) and (1.2), the following two definitions are needed.

<u>Definition 1.1</u>. Let $\bar{z}$ be a vector in E^k and let K^1 be the set of row vectors $z' = (z_0, \bar{z})' \in E^{k+1}$, with the property that there exists at least one $y \in L$ such that $-z_0 + F(y) \leq 0$ and $\underline{f}(y) + \bar{z} \leq 0$, i.e.

$$K^1 = \{z \mid -z_0 + F(y) \leq 0,\ \underline{f}(y) + \bar{z} \leq 0 \text{ for some } y \in L\} .$$

Definition 1.2. Slater's Regularity Condition. There exists at least one $y^* \in L$ such that $\underline{f}(y^*) < 0$.

Under the assumptions that the set K^1 is convex and that Slater's Regularity Condition holds, then for y^o to be an optimal solution of (1.1) it is necessary that y^o and some $x^o \geq 0$ be a saddle point solution of (1.2).

If $L = \{y \mid y \geq 0,\ y \in E^k\}$, i.e. $L = Q_k^+$, and if $\phi(y, x)$ has continuous first derivatives, i.e. ϕ_x and ϕ_y at (y^o, x^o), then the following six "Kuhn-Tucker Conditions" are necessary conditions that the composite vector (y^o, x^o) satisfies (1.2):

$$\phi_x'(y^o, x^o)(x^o) = 0; \tag{1.3}$$

$$\phi_x(y^o, x^o) \leq 0; \tag{1.4}$$

$$x^o \geq 0; \tag{1.5}$$

$$\phi_y'(y^o, x^o)(y^o) - 0; \tag{1.6}$$

$$\phi_y(y^o, x^o) \geq 0; \tag{1.7}$$

and

$$y^o \geq 0 . \tag{1.8}$$

In addition, a sufficient condition to insure that (y^o, x^o) is a saddle point solution is that $\phi(y, x^o)$ is convex in y , i.e.,

$$\phi(y, x^o) \geq \phi(y^o, x^o) + \phi_y'(y^o, x^o)(y - y^o) \quad \text{on} \quad Q_k^+ \ .$$

Figure 1.1 describes the relationship between the general nonlinear minimization problem, the saddle value problem, and the Kuhn-Tucker conditions.

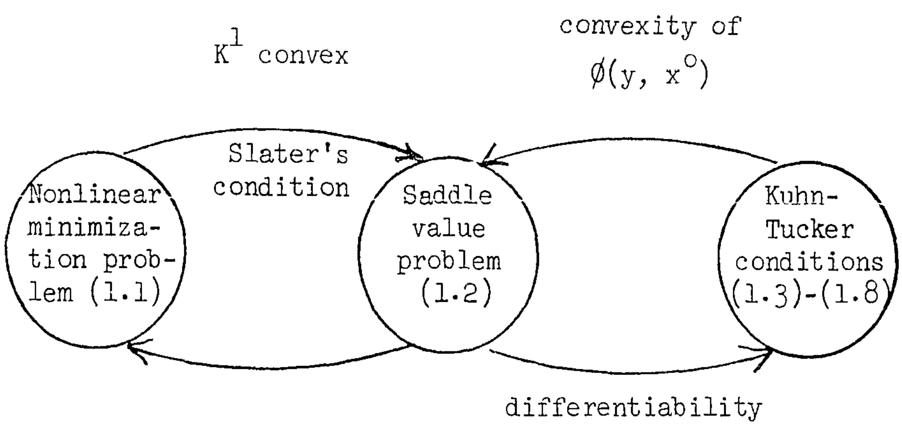

Figure 1.1. Relationship between the general nonlinear minimization problem, the saddle value problem, and the Kuhn-Tucker conditions.

In this study we are concerned with various aspects of quadratic programming. Consequently the objective function $F(y)$ is quadratic and the restrictions $\underline{f}(y)$ are linear in y. Usually the set L is either E^k or Q_k^+. For these situations $F(y)$ is a convex function of y, each component of $\underline{f}(y)$ is a linear function of y, and L is a convex set.

Under these restrictions, as shown by Sposito and David (1971), the set K^1 is convex and Slater's Condition is not needed in establishing the equivalence relationship between (1.1) and (1.2). Especially, if

$F(y)$ is a convex function, and $\underline{f}(y)$ is a componentwise convex function of y, and if L is an arbitrary convex set, then K^1 is convex. Moreover, as underscored by Kunzi and Krelle (1966) and Sposito and David (1971), if the feasible region is polyhedral, then Slater's Regularity Condition is not needed in establishing the fundamental equivalence relationship. Consequently, in this study all of the regularity conditions are automatically satisfied; see Kunzi and Krelle (1966).

Within the class of quadratic programming problems, in Chapter II and III we investigate the problem of least squares subject to inequality restrictions. In particular, in this study we shall extend the concept of closed form least squares solutions to closed form least squares solutions over linear interval constraints

$$P_0 \leq A'\beta \leq P_1$$

as well as linear inequality constraints

$$A'\beta \geq P .$$

Specifically, we are interested in deriving the closed form solution to the problem

$$\begin{aligned} &\text{minimize} \quad (Y-X\beta)'\,(Y-X\beta) \\ &\text{subject to} \quad A'\beta \geq P \qquad \text{or} \qquad P_0 \leq A'\beta \leq P_1 \\ &\qquad\qquad\qquad \beta \in E^k . \end{aligned}$$

The function $F(\beta) = (Y-X\beta)'\,(Y-X\beta)$ of (1.1) is the error sum of squares, which is to be minimized over the constrained region. Since β is unrestricted in E^k, then the six Kuhn-Tucker conditions (1.3)-(1.8) can be reduced to four:

$$\phi_x'(\beta^o, x^o)(x^o) = 0 ; \tag{1.9}$$

$$\phi_x(\beta^o, x^o) \leq 0 ; \tag{1.10}$$

$$x^o \geq 0 ; \tag{1.11}$$

and

$$\phi_\beta(\beta^o, x^o) = 0 . \tag{1.12}$$

In Chapter IV we investigate the bias and mean square error of the closed form estimators developed in Chapters II and III. In particular, it is shown that these biased estimators often have lower mean square error than the unrestricted least squares estimator.

Finally, in Chapter V a generalized upper bounding algorithm is developed to solve a general class of quadratic programming problems. This algorithm reduces the size of the basic matrix so that the total number of computations is greatly reduced in comparison to other standard quadratic programming procedures. This algorithm is extremely useful in solving least squares problems with simple bounds on the parameters.

CHAPTER II. CLOSED FORM SOLUTIONS FOR LEAST SQUARES ESTIMATORS WHEN THE PARAMETERS ARE SUBJECT TO LINEAR INEQUALITY CONSTRAINTS

Introduction

The traditional multiple regression problem involves estimating β in the linear model

$$Y = X\beta + \epsilon ,$$

where Y is a known $n \times 1$ vector, X is a known $n \times k$ matrix, β is a $k \times 1$ vector and ϵ is an $n \times 1$ vector of disturbances such that $\epsilon_i \sim NID(0, \sigma^2)$ for $i = 1, 2, \cdots, n$. Employing least squares as the basis for optimality, the problem to solve becomes minimize the error sum of squares, i.e. $(Y - X\beta)'(Y - X\beta)$. Provided $(X'X)^{-1}$ exists, the best linear unbiased estimator for β is $\hat{\beta} = (X'X)^{-1}X'Y$.

In the above traditional problem, β is completely unrestricted. Often, however, this is not a reasonable assumption, and we have reason to investigate the traditional least squares regression problem subject to constraints of the form $A'\beta \geq P$, where A' is a known $m \times k$ matrix and P is a known $k \times 1$ vector. With the addition of this set of constraints, the traditional regression problem becomes a contrained quadratic programming problem. The problem could be solved iteratively employing algorithms such as those developed by Hildreth (1957) or Wolfe (1957).

Zellner (1961) offers a closed form solution for the simple linear regression case. Thiel and Van de Panne (1960) investigate the general setting and suggest only an iterative method of solution, and later authors discuss inherent difficulties in the type of iteration needed in this procedure. Judge and Takayama (1966) also formulate the above restricted least squares problem as a usual quadratic program. The case when $A' = [1, 0, \ldots, 0]$ and $P = 0$ is examined by Lowell and Prescott (1970) resulting in a closed form solution for this case. They also investigate the bias and mean square error of the constrained estimator. An iterative solution is the result of Mantel's (1969) examination of the general situation, provided the set of constraints contains no superfluous constraints.

Viewing the constrained problem as an application of duality theory, Pyne (1972) successfully derives a closed form solution for β when $A'\beta \geq P$ where A' is $1 \times k$ and P is an arbitrary real number. Pyne shows that his closed form is equivalent to the closed form given by Lowell and Prescott. Subsequent cases are not considered by Pyne. Hence, this chapter will investigate whether it is possible to extend Pyne's initial work.

This study utilizes the Kuhn-Tucker conditions of nonlinear programming which is different than the approach of Pyne. In the next section of this chapter we discuss the Kuhn-Tucker results and matrix theory that direct our method of investigation. We use these results to derive the closed form solution for β when A' is a $1 \times k$ and $2 \times k$ matrix.

We then discuss our findings and difficulties encountered when A' is a matrix of size $m \times k$. We conclude the chapter with a numerical example.

Background and Theory

The least squares problem with inequality conditions can be written as

$$\text{minimize } (Y - X\beta)'(Y - X\beta) = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

$$\text{subject to} \quad A'\beta \geq P$$

$$\beta \quad \text{unrestricted} .$$

We can write the problem equivalently as

$$\text{minimize} \quad \tfrac{1}{2}\beta'X'X\beta - \beta'X'Y$$

$$\text{subject to} \quad P - A'\beta \leq 0$$

$$\beta \quad \text{unrestricted} .$$

The Lagrangian function associated with the saddle value problem is

$$\phi(\beta, \lambda) = \tfrac{1}{2}\beta'X'X\beta - \beta'X'Y + \lambda'(P - A'\beta) .$$

Appealing to the results of Kuhn-Tucker (1951), the following conditions are necessary and sufficient that a feasible solution $(\tilde{\beta}, \lambda^{0})$ solves the saddle value problem:

$$\lambda^{o} \geq 0 \ ; \tag{2.1}$$

$$\phi_{\lambda}(\tilde{\beta}, \lambda^{o}) \leq 0 \qquad \text{or} \qquad (P - A'\tilde{\beta}) \leq 0 \ ; \tag{2.2}$$

$$(\lambda^{o'})\phi'_{\lambda}(\tilde{\beta}, \lambda^{o}) = 0 \qquad \text{or} \qquad \lambda^{o'}(P - A'\tilde{\beta}) = 0 \ ; \tag{2.3}$$

and

$$\phi_{\beta}(\tilde{\beta}, \lambda^{o}) = 0 \qquad \text{or} \qquad X'X\tilde{\beta} - X'Y - A\lambda^{o} = 0 \ . \tag{2.4}$$

There are only four conditions instead of the classical six due to the fact that β is unrestricted and that the objective function, $F(y) = \frac{1}{2}\beta'X'X\beta - \beta'X'Y$ is convex, $\frac{1}{2}\beta'X'X\beta$ is a quadratic, convex function of β and $-\beta'X'Y$ is linear. The partial derivatives of $\phi(\beta, \lambda)$ with respect to β can be set equal to zero, resulting in condition (2.4) above.

In the development of the rest of this chapter, the following two lemmas will prove fundamental.

Lemma 2.1. If $(X'X)$ is a positive definite matrix, then $(X'X)^{-1}$ is positive definite.

Proof:

Since $X'X$ is positive definite, then $\det (X'X) \neq 0$ and so $(X'X)^{-1}$ exists. Given $y \neq 0$, it follows that $\bar{y} = (X'X)^{-1}y \neq 0$, so that $\bar{y}'(X'X)\bar{y} > 0$. Therefore $y'(X'X)^{-1}y = \bar{y}'(X'X)(X'X)^{-1}(X'X)\bar{y} = \bar{y}'(X'X)\bar{y} > 0$, so that $(X'X)^{-1}$ is positive definite. $\square$

Lemma 2.2. If $(X'X)^{-1}$ is a positive definite matrix of size $k \times k$ and if A' is a matrix with m rows and k columns, that is of full row rank, then $A'(X'X)^{-1}A$ is a symmetric, positive definite matrix.

Proof:

Given any vector $y \neq 0$, let $\bar{y} = Ay$. Since A' and A are of rank m, then $Ay = \bar{y} \neq 0$. Therefore, $y'(A'(X'X)^{-1}A)y = (Ay)'(X'X)^{-1}(Ay) = \bar{y}'(X'X)^{-1}\bar{y} > 0$, which implies that $A'(X'X)^{-1}A$ is positive definite.

Moreover, $A'(X'X)^{-1}A$ is symmetric since $(X'X)^{-1}$ is symmetric. Therefore, the result follows. $\square$

Corollary 2.1. If A' is not of full row rank, then $A'(X'X)^{-1}A$ is positive semi-definite.

From now on, we will assume that $(X'X)^{-1}$ exists and that A' is of full row rank, so that $A'(X'X)^{-1}A$ is positive definite.

Hildreth (1957) developed an asymptotic algorithm to solve the general problem

$$\text{minimize} \quad \{p'x + x'Cx \mid Ax \leq b\} , \tag{2.5}$$

where A is an $m \times k$ matrix, C is a positive definite $k \times k$ matrix and, hence, the objective function is necessarily strictly convex. In particular, by appealing to the four Kuhn-Tucker conditions (2.1), (2.2), (2.3) and (2.4), a nonclassical dual quadratic problem can be derived which also satisfies the Kuhn-Tucker conditions.

To show this we introduce slack variables y, so that the four Kuhn-Tucker conditions become:

$$Ax + y = b \ ; \tag{2.6}$$

$$2Cx + A'\lambda = -p \ ; \tag{2.7}$$

$$\lambda \geq 0 \ , \quad y \geq 0 \ ; \tag{2.8}$$

and

$$\lambda' y = 0 \ . \tag{2.9}$$

Solving for x in (2.7), we have that

$$x = x(\lambda) = -\tfrac{1}{2}C^{-1}(A'\lambda + p) \ .$$

If we define $h = \frac{1}{2}AC^{-1}p + b$ and $G = \frac{1}{4} AC^{-1}A'$, and substitute the value of x into (2.6), then the four conditions can be reduced to:

$$2G\lambda - y = -h \ ; \tag{2.10}$$

$$\lambda \geq 0 \ , \quad y \geq 0 \ ; \tag{2.11}$$

and

$$\lambda' y = 0 \ . \tag{2.12}$$

From Lemma 2.2, G is positive definite or semi-definite depending on the rank of A. Consequently, (2.10), (2.11) and (2.12) are exactly the Kuhn-Tucker conditions of the dual quadratic problem

$$\text{minimize} \quad \{h'\lambda + \lambda' G\lambda \mid \lambda \geq 0\} \; . \tag{2.13}$$

Since the Kuhn-Tucker conditions are necessary and sufficient for both (2.5) and (2.13), we have the following lemma:

<u>Lemma 2.3</u>. $x^o = x^o(\lambda^o)$ is an optimal solution of (2.5) if and only if λ^o is an optimal solution of the dual problem (2.13).

Hildreth does not touch on the topic of closed form expressions for the original problem, but only presents an iterative procedure to solve the simpler nonclassical dual problem.

A' is a 1 x k Matrix

In order to solve the problem

$$\text{minimize} \quad \tfrac{1}{2}\beta' X' X\beta - \beta' X' Y$$

$$\text{subject to} \quad P - A'\beta \leq 0$$

$$\beta \quad \text{unrestricted}$$

we must find $(\tilde{\beta}, \lambda^o)$ that satisfies the following Kuhn-Tucker conditions:

$$\lambda^o \geq 0 \; ; \tag{2.1}$$

$$P - A'\tilde{\beta} \leq 0 \; ; \tag{2.2}$$

$$\lambda^{o'}(P - A'\tilde{\beta}) = 0 \; ; \tag{2.3}$$

and

$$X'X\tilde{\beta} - X'Y - A\lambda^{o} = 0 \; . \tag{2.4}$$

Solving for $\tilde{\beta}$ in (2.4) we see that

$$\begin{aligned}\tilde{\beta} &= (X'X)^{-1}\,[X'Y + A\lambda^{o}] \\ &= (X'X)^{-1}X'Y + (X'X)^{-1}A\lambda^{o} \\ &= \hat{\beta} + (X'X)^{-1}A\lambda^{o} \; ,\end{aligned} \tag{2.14}$$

where $\hat{\beta}$ is the best linear unbiased estimator of β for the traditional unrestricted least squares problem. Consequently, to find $\tilde{\beta}$, we first determine $\hat{\beta}$ and then adjust $\hat{\beta}$ to $\tilde{\beta}$ so that $\tilde{\beta}$ lies within the constrained region.

<u>Case 1.</u> <u>$\hat{\beta}$ satisfies the restriction, i.e. $P - A'\hat{\beta} \leq 0$</u>

$(\tilde{\beta}, \lambda^{o}) = (\hat{\beta}, 0)$ satisfies the conditions (2.1)-(2.4), so that $\tilde{\beta} = \hat{\beta} = (X'X)^{-1}X'Y$.

<u>Case 2.</u> <u>$\hat{\beta}$ violates the restriction, i.e. $P - A'\hat{\beta} > 0$</u>

From condition (2.3) either λ^{o} or $P - A'\tilde{\beta}$ equals zero, or both.

<u>Subcase 1</u> $\lambda^{o} = 0$.

From condition (2.2) and equation (2.14),

$$P - A'\tilde{\beta} = P - A'\hat{\beta} - A'(X'X)^{-1}A\lambda^{o} \leq 0 \; .$$

However, since $\lambda^{o} = 0$ and $P - A'\hat{\beta} > 0$, this subcase is impossible.

Subcase 2 $\lambda^o > 0$.

From condition (2.3) and equation (2.14)

$$P - A'\tilde{\beta} = P - A'\hat{\beta} - A'(X'X)^{-1}A\lambda^o = 0 . \tag{2.15}$$

Solving for λ^o in (2.15),

$$\lambda^o = \frac{P - A'\hat{\beta}}{A'(X'X)^{-1}A} . \tag{2.16}$$

Substituting for λ^o in (2.14), we have that

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}A \left(\frac{P - A'\hat{\beta}}{A'(X'X)^{-1}A} \right) . \tag{2.17}$$

Table 2.1 summarizes the above discussion.

Table 2.1. Constrained least squares solution for β where $A'\beta \geq P$ and A' is a $1 \times k$ matrix

$A'\hat{\beta} \geq P$	$A'\hat{\beta} < P$
$\hat{\beta}$	$\hat{\beta} + (X'X)^{-1}A\,[A'(X'X)^{-1}A]^{-1}(P - A'\hat{\beta})$

A' is a $2 \times k$ Matrix

Following the same procedure as in the previous section, we will derive the closed form solution for β by adjusting $\hat{\beta}$ to $\tilde{\beta}$ so that $\tilde{\beta}$ lies in the constrained region. Since it is possible for $\hat{\beta}$ to satisfy none, one, or both of the given linear restrictions, it will prove convenient to partition the matrices A and A' and vectors P and λ as follows:

$$A = (A_1, A_2) \; ; \qquad A' = \begin{pmatrix} A_1' \\ A_2' \end{pmatrix} \; ; \qquad P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix};$$

and

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} .$$

We define $D = \begin{pmatrix} A_1' \\ A_2' \end{pmatrix} (X'X)^{-1} (A_1, A_2)$; D is a 2×2 matrix and is the variance-covariance matrix of $A_1'\hat{\beta}$ and $A_2'\hat{\beta}$. From Lemma 2.2, D is a symmetric, positive definite matrix, so that $\det D > 0$ and D^{-1} exists. Consequently, if we define the components of D as

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} ,$$

where d_{11} and $d_{22} > 0$, then

$$D^{-1} = \frac{1}{\det D} \begin{pmatrix} d_{22} & -d_{12} \\ -d_{12} & d_{11} \end{pmatrix} .$$

$$\text{Det } D = d_{22}\, d_{11} - d_{12}^2 .$$

In partitioned form, the Kuhn-Tucker conditions (2.1)-(2.3) become:

$$\lambda_1^o \geq 0 , \quad \lambda_2^o \geq 0 ; \tag{2.1}$$

$$P_1 - A_1'\tilde{\beta} \leq 0 , \quad P_2 - A_2'\tilde{\beta} \leq 0 ; \tag{2.2}$$

and

$$(\lambda_1^o, \lambda_2^o) \begin{pmatrix} P_1 - A_1'\tilde{\beta} \\ P_2 - A_2'\tilde{\beta} \end{pmatrix}$$

$$= \lambda_1^o(P_1 - A_1'\tilde{\beta}) + \lambda_2^o(P_2 - A_2'\tilde{\beta}) = 0 . \tag{2.3}$$

From (2.1) and (2.2), $\lambda_i^o(P_i - A_i'\tilde{\beta}) \leq 0$ for $i = 1, 2,$ so that (2.3) holds only if

$$\lambda_i^o(P_i - A_i'\tilde{\beta}) = 0 \quad \text{for} \quad i = 1, 2 . \tag{2.18}$$

Equality (2.18) implies that either $\lambda_i^o = 0$ or $P_i - A_i'\tilde{\beta} = 0$ for $i = 1, 2$. Both of each pair could be zero.

Once again from the Kuhn-Tucker condition (2.4) we can solve for $\tilde{\beta}$ in terms of $\hat{\beta}$.

$$X'X\tilde{\beta} - X'Y - (A_1,\ A_2)\begin{pmatrix}\lambda_1^o \\ \lambda_2^o\end{pmatrix} = 0 \ . \tag{2.4}$$

Consequently,

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}(A_1,\ A_2)\begin{pmatrix}\lambda_1^o \\ \lambda_2^o\end{pmatrix} \tag{2.14}$$

where

$$\hat{\beta} = (X'X)^{-1}\ X'Y \ .$$

Substituting for $\tilde{\beta}$ in condition (2.2),

$$\begin{pmatrix} P_1 - A_1'\tilde{\beta} \\ P_2 - A_2'\tilde{\beta}\end{pmatrix} = \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta}\end{pmatrix} - \begin{pmatrix} A_1' \\ A_2'\end{pmatrix}(X'X)^{-1}(A_1,\ A_2)\begin{pmatrix}\lambda_1^o \\ \lambda_2^o\end{pmatrix} \leq \begin{pmatrix}0 \\ 0\end{pmatrix} . \tag{2.19}$$

Rewriting (2.19),

$$D\begin{pmatrix}\lambda_1^o \\ \lambda_2^o\end{pmatrix} \geq \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta}\end{pmatrix} \equiv \begin{pmatrix} r_1 \\ r_2\end{pmatrix} . \tag{2.20}$$

<u>Case 1.</u> <u>$\hat{\beta}$ satisfies both restrictions, i.e., $\begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$</u>

$$\left(\tilde{\beta}, \begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix}\right) = \left(\hat{\beta}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)$$ satisfies the conditions (2.1)-(2.4), so that $\tilde{\beta} = \hat{\beta} = (X'X)^{-1} X'Y$.

<u>Case 2.</u> <u>$\hat{\beta}$ violates both of the restrictions, i.e. , $\begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \end{pmatrix}$</u>

From (2.18) either λ_i^o , $P_i - A_i'\tilde{\beta}$, or both equal zero for $i = 1, 2$.

<u>Subcase 1</u> $\lambda_1^o = 0$; $\lambda_2^o = 0$.

This combination of λ_i^o is not sufficient as it does not satisfy inequality (2.20).

$$D \begin{pmatrix} 0 \\ 0 \end{pmatrix} \not\geq \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \quad \text{since} \quad \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

<u>Subcase 2</u> $\lambda_1^o = 0$; $\lambda_2^o > 0$.

Since $\lambda_2^o > 0$, then from (2.18), $P_2 - A_2'\tilde{\beta} = 0$. Consequently (2.20) becomes

$$D \begin{pmatrix} 0 \\ \lambda_2^o \end{pmatrix} \begin{matrix} \geq \\ = \end{matrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}, \quad \text{or}$$

$$\begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} 0 \\ \lambda_2^o \end{pmatrix} \begin{matrix} \geq \\ = \end{matrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} . \tag{2.21}$$

Expression (2.21) is consistent only if

$$d_{12}\lambda_2^o \geq r_1 \qquad \text{and} \qquad d_{22}\lambda_2^o = r_2 . \tag{2.22}$$

If $d_{12} \leq 0$, then (2.22) implies that $d_{12} \geq \frac{r_1}{\lambda_2^o}$. This inequality can never hold, since $\frac{r_1^+}{\lambda_2^{o+}}$ is always positive and d_{12} is assumed to be nonpositive.

If $d_{12} > 0$, then from (2.22) $\lambda_2^o \geq \frac{r_1^+}{d_{12}^+}$ and $\lambda_2^o = \frac{r_2^+}{d_{22}^+}$. Together these imply that

$$\frac{r_2}{d_{22}} \geq \frac{r_1}{d_{12}} \quad \text{or} \quad \frac{r_1}{r_2} \leq \frac{d_{12}}{d_{22}} . \tag{2.23}$$

Therefore, whenever $d_{12} > 0$ and $\frac{r_1}{r_2} \leq \frac{d_{12}}{d_{22}}$ then the closed form solution for β is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} (A_1, A_2) \begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1}(A_1, A_2) \begin{pmatrix} 0 \\ \frac{r_2}{d_{22}} \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{r_2}{d_{22}} \right) .$$

<u>Subcase 3</u> $\lambda_1^o > 0$; $\lambda_2^o = 0$.

Since $\lambda_1^o > 0$, then from (2.18), $P_1 - A_1'\tilde{\beta} = 0$. Therefore (2.20) becomes

$$\begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ 0 \end{pmatrix} \begin{matrix} = \\ \geq \end{matrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} . \tag{2.24}$$

Expression (2.24) is consistent only if

$$d_{11}\lambda_1^o = r_1 \quad \text{and} \quad d_{12}\lambda_1^o \geq r_2 . \tag{2.25}$$

Putting both expressions of (2.25) together, then $\frac{r_1}{d_{11}} d_{12} \geq r_2$.

If $d_{12} > 0$, then $\frac{r_1^+}{d_{11}^+} \geq \frac{r_2^+}{d_{12}^+}$ or $\frac{r_1}{r_2} \geq \frac{d_{11}}{d_{12}}$.

If $d_{12} \leq 0$, then the expression $\frac{r_1^+}{d_{11}^+} d_{12}^- \geq r_2^+$ can never hold.

Therefore, if $d_{12} > 0$ and $\frac{r_1}{r_2} \geq \frac{d_{11}}{d_{12}}$ then the closed form solution for β is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_1 \left(\frac{r_1}{d_{11}} \right) .$$

Subcase 4 $\quad \lambda_1^o > 0 \; ; \; \lambda_2^o > 0$.

Since both λ_1^o and $\lambda_2^o > 0$. then from (2.18) $P_1 - A_1'\tilde{\beta} = 0$ and $P_2 - A_2'\tilde{\beta} = 0$. Therefore inequality (2.20) becomes an equality, i.e.

$$\begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} . \tag{2.26}$$

Solving for λ_1^o and λ_2^o in (2.26),

$$\begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix} = \frac{1}{\det D} \begin{pmatrix} d_{22} & -d_{12} \\ -d_{12} & d_{11} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} .$$

Hence

$$\lambda_1^o = \frac{d_{22}r_1 - d_{12}r_2}{\det D} \; ; \quad \lambda_2^o = \frac{d_{11}r_2 - d_{12}r_1}{\det D} . \tag{2.27}$$

Since λ_1^o, λ_2^o , and $\det D > 0$. then (2.27) holds only if

$$\overset{+}{d}_{22}\overset{+}{r}_1 - d_{12}\overset{+}{r}_2 > 0 \quad \text{and} \quad \overset{+}{d}_{11}\overset{+}{r}_2 - d_{12}\overset{+}{r}_1 > 0 . \tag{2.28}$$

If $d_{12} \leq 0$, then (2.28) is always true. If $d_{12} > 0$ then (2.28) holds whenever $d_{11}r_2 > d_{12}r_1$ and $d_{22}r_1 > d_{12}r_2$, which together imply that

$$\frac{d_{12}}{d_{22}} < \frac{r_1}{r_2} < \frac{d_{11}}{d_{12}} .$$

Consequently, if $d_{12} < 0$ or if $d_{12} > 0$ and $\frac{d_{12}}{d_{22}} < \frac{r_1}{r_2} < \frac{d_{11}}{d_{12}}$ then the closed form solution for β is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}(A_1, A_2) \begin{pmatrix} \dfrac{d_{22}r_1 - d_{12}r_2}{\det D} \\ \\ \dfrac{d_{11}r_2 - d_{12}r_1}{\det D} \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1} AD^{-1} (P - A'\hat{\beta}) .$$

<u>Summary of Case 2</u> Both constraints are violated.

a) If $\operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} \leq 0$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} AD^{-1} (P - A'\tilde{\beta})$.

b) If $\operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} > 0$

and

i) $$\frac{d_{12}}{d_{22}} < \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} < \frac{d_{11}}{d_{12}}$$

then $\quad \tilde{\beta} = \hat{\beta} + (X'X)^{-1}A\, D^{-1}(P - A'\hat{\beta})$;

or

$$\text{ii)} \quad \frac{d_{12}}{d_{22}} \geq \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$$

then $\quad \tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\dfrac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$;

or if

$$\text{iii)} \quad \frac{d_{11}}{d_{12}} \leq \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$$

then $\quad \tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_1 \left(\dfrac{P_1 - A_1'\hat{\beta}}{d_{11}} \right)$.

Lemma 2.4. Closed forms b(i) and b(ii) above are equivalent when

$$\frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} = \frac{d_{12}}{d_{22}} .$$

Proof:

Closed form b(i) is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A[A'(X'X)^{-1}A]^{-1} [P - A'\hat{\beta}]$$

$$= \hat{\beta} + (X'X)^{-1} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{k1} & a_{k2} \end{pmatrix} \frac{1}{\det D} \begin{pmatrix} d_{22} & -d_{12} \\ -d_{12} & d_{11} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} ,$$

since $r_1 = P_1 - A_1'\hat{\beta}$, $r_2 = P_2 - A_2'\hat{\beta}$, and $D = A'(X'X)^{-1}A$,

$$= \hat{\beta} + (X'X)^{-1} \frac{1}{\det D} \begin{pmatrix} a_{11}d_{22} - d_{12}a_{12} \; , & -a_{11}d_{12} + a_{12}d_{11} \\ a_{21}d_{22} - d_{12}a_{22} \; , & -a_{21}d_{12} + a_{22}d_{11} \\ \vdots \quad \vdots & \vdots \\ a_{k1}d_{22} - d_{12}a_{k2} \; , & -a_{k1}d_{12} + a_{k2}d_{11} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1} \frac{1}{\det D} \begin{pmatrix} (a_{11}d_{22} - a_{12}d_{12})r_1 + r_2(a_{12}d_{11} - a_{11}d_{12}) \\ (a_{21}d_{22} - a_{22}d_{12})r_1 + r_2(a_{22}d_{11} - a_{21}d_{12}) \\ \vdots \qquad \vdots \qquad \vdots \\ (a_{k1}d_{22} - a_{k2}d_{12})r_1 + r_2(a_{k2}d_{11} - a_{k1}d_{12}) \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1} \frac{1}{\det D} \begin{pmatrix} a_{11}d_{22}r_1 - a_{12}d_{12}r_1 + a_{12}d_{11}r_2 - a_{11}d_{12}r_2 \\ a_{21}d_{22}r_1 - a_{22}d_{12}r_1 + a_{22}d_{11}r_2 - a_{21}d_{12}r_2 \\ \vdots \qquad \vdots \qquad \vdots \\ a_{k1}d_{22}r_1 - a_{k2}d_{12}r_1 + a_{k2}d_{11}r_2 - a_{k1}d_{12}r_2 \end{pmatrix} . \tag{2.29}$$

When $\frac{r_1}{r_2} = \frac{d_{12}}{d_{22}}$ then $a_{11}d_{22}r_1 = a_{11}d_{12}r_2$ so that in (2.29) the first and last terms of each row drop. The vector on the right of (2.29) becomes

$$\begin{pmatrix} a_{12}d_{11}r_2 - a_{12}d_{12}r_1 \\ a_{22}d_{11}r_2 - a_{22}d_{12}r_1 \\ \vdots \\ a_{k2}d_{11}r_2 - a_{k2}d_{12}r_1 \end{pmatrix} = \begin{pmatrix} a_{12}(d_{11}r_2 - d_{12}r_1) \\ a_{22}(d_{11}r_2 - d_{12}r_1) \\ \vdots \\ a_{k2}(d_{11}r_2 - d_{12}r_1) \end{pmatrix}$$

$$= \begin{pmatrix} a_{12}(d_{11}r_2 - d_{12} \cdot \frac{d_{12}r_2}{d_{22}}) \\ a_{22}(d_{11}r_2 - d_{12} \cdot \frac{d_{12}r_2}{d_{22}}) \\ \vdots \\ a_{k2}(d_{11}r_2 - d_{12} \cdot \frac{d_{12}r_2}{d_{22}}) \end{pmatrix}$$

$$= (d_{11}r_2 - d_{12} \cdot \frac{d_{12}r_2}{d_{22}}) \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix}. \tag{2.30}$$

Multiplying (2.30) by $\frac{1}{\det D}$,

$$\left(\frac{d_{11}r_2 - d_{12} \cdot \frac{d_{12}r_2}{d_{22}}}{d_{11}d_{22} - d_{12}^2} \right) \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix} = \frac{\frac{r_2}{d_{22}}(d_{11}d_{22} - d_{12}^2)}{(d_{11}d_{22} - d_{12}^2)} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix}$$

$$= A_2\left(\frac{r_2}{d_{22}}\right) .$$

Equation (2.29) is then $\hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}}\right)$ which is the closed from b(ii) . ▯

Lemma 2.5. Closed forms b(i) and b(iii) are equivalent when

$$\frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} = \frac{d_{11}}{d_{12}} .$$

Proof:

The proof is the same as that of Lemma 2.4. ▯

Case 3. $\hat{\beta}$ violates the second restriction, i.e., $\begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \begin{matrix} \leq \\ > \end{matrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

We must adjust $\hat{\beta}$ in order to satisfy the second restriction, however, we must simultaneously guarantee that the first restriction continues to be satisfied.

Subcase 1 $\lambda^o = 0$; $\lambda_2^o = 0$.

Once again this combination of λ_i^o is not sufficient as it does not satisfy (2.20).

$$D \begin{pmatrix} 0 \\ 0 \end{pmatrix} \not\geq \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \quad \text{since} \quad \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \begin{matrix} \leq \\ > \end{matrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

<u>Subcase 2</u> $\lambda_1^o = 0$; $\lambda_2^o > 0$.

Since $\lambda_2^o > 0$, then from (2.18) $P_2 - A_2'\tilde{\beta} = 0$. Therefore, from (2.20), λ_2^o must satisfy

$$\begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} 0 \\ \lambda_2^o \end{pmatrix} \begin{matrix} \geq \\ = \end{matrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} . \tag{2.31}$$

Expression (2.31) is consistent only if $d_{12}\lambda_2^{o+} \geq r_1$ and $\lambda_2^{o+} d_{22}^{+} = r_2^{+}$, which together imply that

$$\frac{d_{12}}{d_{22}^{+}} \geq \frac{r_1}{r_2^{+}} . \tag{2.32}$$

If $d_{12} \geq 0$, then inequality (2.32) always holds since r_1 is nonpositive.

If $d_{12} < 0$, and if $\frac{d_{12}}{d_{22}} \geq \frac{r_1}{r_2}$ then inequality (2.32) will be satisfied and λ^o will be positive.

Therefore, whenever $d_{12} \geq 0$ or $d_{12} < 0$ and $\frac{d_{12}}{d_{22}} \geq \frac{r_1}{r_2}$ the closed form solution for β is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} (A_1, A_2) \begin{pmatrix} 0 \\ \frac{r_2}{d_{22}} \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{r_2}{d_{22}} \right) .$$

<u>Subcase 3</u> $\lambda_1^o > 0$; $\lambda_2^o = 0$.

Since $\lambda_1^o > 0$, then (2.18) implies that $P_1 - A_1'\tilde{\beta} = 0$. This combination of λ_i^o is not sufficient since (2.20) does not hold, i.e.

$$\begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ 0 \end{pmatrix} \begin{matrix} \neq \\ \geq \end{matrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

since d_{11} and $\lambda_1^o > 0$ and r_1 is nonpositive.

<u>Subcase 4</u> $\lambda_1^o > 0$; $\lambda_2^o > 0$.

Because $\lambda_i^o > 0$ for $i = 1, 2$, then from (2.18) $P_i - A_i'\tilde{\beta} = 0$ for $i = 1, 2$. Therefore, we can solve for λ_1^o and λ_2^o using equation (2.20).

$$\begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} ,$$

so that

$$\begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix} = \frac{1}{\det D} \begin{pmatrix} d_{22}r_1 - d_{12}r_2 \\ d_{11}r_2 - d_{12}r_1 \end{pmatrix} . \tag{2.33}$$

Now, since λ_1^o, λ_2^o , and $\det D > 0$ we have that (2.33) will hold only if

$$d_{22}^{+}r_1 - d_{12}r_2^{+} > 0 \quad \text{and} \quad d_{11}^{+}r_2^{+} - d_{12}r_1 > 0 \; . \tag{2.34}$$

If $d_{12} < 0$, then since r_1 is nonpositive, (2.34) will hold if $d_{22}r_1 > d_{12}r_2$ and $d_{11}r_2 > d_{12}r_1$, or $\frac{r_1}{r_2} > \frac{d_{12}}{d_{22}}$ and $\frac{r_1}{r_2} > \frac{d_{11}}{d_{12}}$. Since $\det D > 0$ and $d_{12} < 0$, then $\frac{d_{12}}{d_{22}} < \frac{r_1}{r_2}$ implies $\frac{d_{11}}{d_{12}} < \frac{r_1}{r_2}$.

If $d_{12} \geq 0$, then since r_1 is nonpositive, (2.34) can never hold.

Therefore, if $d_{12} < 0$ and $\frac{d_{12}}{d_{22}} < \frac{r_1}{r_2}$, then the closed form solution for β is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} (A_1, A_2) \begin{pmatrix} \dfrac{d_{22}r_1 - d_{12}r_2}{\det D} \\ \\ \dfrac{d_{11}r_2 - d_{12}r_1}{\det D} \end{pmatrix}$$

$$= \hat{\beta} + (X'X)^{-1} A D^{-1} (P - A'\hat{\beta}) \; .$$

<u>Summary of Case 3</u> $\quad \hat{\beta}$ violates the second constraint.

a) If $\operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} < 0$

and

i) $\dfrac{d_{12}}{d_{22}} \geq \dfrac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\dfrac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$;

or

$$\text{ii)} \quad \frac{d_{12}}{d_{22}} < \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A D^{-1} (P - A'\hat{\beta})$.

b) If $\text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} \geq 0$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\dfrac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$.

<u>Case 4.</u> <u>$\hat{\beta}$ violates the first constraint, i.e., $\begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \begin{matrix} > \\ \leq \end{matrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$</u>

If we use the transformation $A_1'^{*} = A_2'$, $P_1^{*} = P_2$, $P_2^{*} = P_1$, and $A_2'^{*} = A_1'$, then this case becomes

$$\begin{pmatrix} P_1^{*} - A_1'^{*}\hat{\beta} \\ P_2^{*} - A_2'^{*}\hat{\beta} \end{pmatrix} \begin{matrix} \geq \\ > \end{matrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which is Case 3.

<u>Summary of Case 4</u> $\hat{\beta}$ violates the first constraint.

a) If $\text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} < 0$

and

$$\text{i)} \quad \frac{d_{12}}{d_{11}} \geq \frac{P_2 - A_2'\hat{\beta}}{P_1 - A_1'\hat{\beta}}$$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_1 \left(\frac{P_1 - A_1'\hat{\beta}}{d_{11}} \right)$;

ii) $\frac{d_{12}}{d_{11}} < \frac{P_2 - A_2'\hat{\beta}}{P_1 - A_1'\hat{\beta}}$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A D^{-1} (P - A'\hat{\beta})$.

b) If $\text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} \geq 0$

then $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_1 \left(\frac{P_1 - A_1'\hat{\beta}}{d_{11}} \right)$.

Tables 2.2, 2.3 and 2.4 summarize the discussion of this section. In these tables $A = (A_1, A_2)$, $P - A'\hat{\beta} = \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$, and $D = A'(X'X)^{-1}A = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix}$.

A' is an m x k Matrix

The m consistent linear restrictions on β are $A'\beta \geq P$. If $\hat{\beta}$ violates all of these restrictions then we have the following lemma:

<u>Lemma 2.6</u>. Whenever all restrictions are violated by $\hat{\beta}$ and $\lambda^o = D^{-1}(P - A'\hat{\beta}) > 0$, then the least squares solution for β is $\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A D^{-1}(P - A'\hat{\beta})$, where $D = A'(X'X)^{-1}A$.

Proof:

We must show that $(\tilde{\beta}, \lambda^o)$ satisfies the Kuhn-Tucker conditions

Table 2.2. Constrained least squares solution for β where $A'\beta \geq P$, A' is a $2 \times k$ matrix, and $A'\hat{\beta} < P$

Condition	$d_{12} \leq 0$	$d_{12} > 0$
none	$\hat{\beta} + (X'X)^{-1} A D^{-1}(P - A'\hat{\beta})$	
$\frac{d_{12}}{d_{22}} < \frac{r_1}{r_2} < \frac{d_{11}}{d_{12}}$		$\hat{\beta} + (X'X)^{-1} A D^{-1} (P - A'\hat{\beta})$
$\frac{d_{12}}{d_{22}} \geq \frac{r_1}{r_2}$		$\hat{\beta} + (X'X)^{-1} A_2 \left(\frac{r_2}{d_{22}} \right)$
$\frac{d_{11}}{d_{12}} \leq \frac{r_1}{r_2}$		$\hat{\beta} + (X'X)^{-1} A_1 \left(\frac{r_1}{d_{11}} \right)$

Table 2.3. Constrained least squares solution for β where $A'\beta \geq P$, A' is a $2 \times k$ matrix, and $A_1'\hat{\beta} < P_1$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
none		$\hat{\beta} + (X'X)^{-1} A_1 (\frac{r_1}{d_{11}})$
$\frac{d_{12}}{d_{11}} \geq \frac{r_2}{r_1}$	$\hat{\beta} + (X'X)^{-1} A_1 (\frac{r_1}{d_{11}})$	
$\frac{d_{12}}{d_{11}} < \frac{r_2}{r_1}$	$\hat{\beta} + (X'X)^{-1} A D^{-1} (P - A'\hat{\beta})$	

Table 2.4. Constrained least squares solution for β where $A'\beta \geq P$, A' is a $2 \times k$ matrix, and $A_2'\hat{\beta} < P_2$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
none		$\hat{\beta} + (X'X)^{-1} A_2 \left(\frac{r_2}{d_{22}} \right)$
$\frac{d_{12}}{d_{22}} \geq \frac{r_1}{r_2}$	$\hat{\beta} + (X'X)^{-1} A_2 \left(\frac{r_2}{d_{22}} \right)$	
$\frac{d_{12}}{d_{22}} < \frac{r_1}{r_2}$	$\hat{\beta} + (X'X)^{-1} A D^{-1} (P - A'\hat{\beta})$	

(2.1)-(2.4).

(2.1): $\lambda^{o} = D^{-1}(P - A'\hat{\beta})$, which is > 0 .

(2.2): $P - A'\tilde{\beta} = (P - A'\hat{\beta}) - [A'(X'X)^{-1}AD^{-1}(P - A'\hat{\beta})]$

$= (P - A'\hat{\beta}) - [DD^{-1}(P - A'\hat{\beta})]$

$= P - A'\hat{\beta} - (P - A'\hat{\beta}) = 0$, which is ≤ 0 .

(2.3): $\lambda^{o'}(P - A'\tilde{\beta}) = \lambda^{o'} \cdot 0 = 0$.

(2.4): $(X'X)\tilde{\beta} - X'Y - A\lambda^{o}$

$= (X'X)[\hat{\beta} + (X'X)^{-1}AD^{-1}(P - A'\hat{\beta})] - X'Y - AD^{-1}(P - A'\hat{\beta})$

$= (X'X)\hat{\beta} + (X'X)(X'X)^{-1}AD^{-1}(P-A'\hat{\beta}) - X'Y - AD^{-1}(P-A'\hat{\beta})$

$= (X'X)(X'X)^{-1}(X'Y) + AD^{-1}(P - A'\hat{\beta}) - X'Y - AD^{-1}(P-A'\hat{\beta})$

$= 0$.

Since all four Kuhn-Tucker conditions hold, the result of the lemma follows. $\square$

When all restrictions are violated by $\hat{\beta}$, then $P - A'\hat{\beta} > 0$. Therefore, a sufficient condition for $\lambda^{o} = D^{-1}(P - A'\hat{\beta})$ to be positive is that all elements of D^{-1} are nonnegative with at least one element strictly positive. We will derive conditions on D in order to find a sufficient condition for λ^{o} to be positive.

Case 1. m = 3

$D = A'(X'X)^{-1}A$ is a 3×3 positive definite matrix. Therefore

$$D = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}$$

where a,b,c, and $\det D > 0$ and d,e, and f are unrestricted. Using the formula for matrix inversion by bordering from page 71 of Hemmerle (1967),

$$D^{-1} = (\bar{d}_{ij}) ,$$

where

$$\bar{d}_{11} = \frac{(be - df)^2}{\alpha(\det D)^2} + \frac{b}{\det D} , \tag{2.35}$$

$$\bar{d}_{12} = \frac{(be - fd)(af - ed)}{\alpha(\det D)^2} - \frac{d}{\det D} , \tag{2.36}$$

$$\bar{d}_{13} = \frac{(fd - be)}{\alpha(\det D)} , \tag{2.37}$$

$$\bar{d}_{21} = \bar{d}_{12} ,$$

$$\bar{d}_{22} = \frac{(af - ed)^2}{\alpha(\det D)^2} + \frac{a}{\det D} \tag{2.38}$$

$$\bar{d}_{23} = \frac{ed - af}{\alpha(\det D)} , \tag{2.39}$$

$$\bar{d}_{31} = \bar{d}_{13} ,$$

$$\bar{d}_{32} = \bar{d}_{23} ,$$

and

$$\bar{d}_{33} = \frac{1}{\alpha} ,$$

where

$$\alpha = c(\det D) - be^2 - af^2 + 2def . \tag{2.40}$$

Since D is positive definite, then D^{-1} is also positive definite, so that $\bar{d}_{11}$, $\bar{d}_{22}$, $\bar{d}_{33}$, and $\det D$ are necessarily all greater than zero. Due to the positivity of $\bar{d}_{33}$, α is therefore positive.

From (2.3.7), $\bar{d}_{13}$ and $\bar{d}_{31}$ are nonnegative when $(fd - be) \geq 0$. Likewise from (2.39) $\bar{d}_{23}$ and $\bar{d}_{32}$ are nonnegative when $(ed - af) \geq 0$. Consequently the first term in (2.36) is nonnegative whenever $(fd - be)$ and $(ed - af)$ are nonnegative. Therefore a sufficient condition that $\bar{d}_{12}$ and $\bar{d}_{21}$ are nonnegative is that $-d$ is nonnegative, i.e., $d \leq 0$.

In summary, therefore, whenever all constraints are violated, the following conditions are sufficient conditions to insure that $\lambda^{o} > 0$:

$$fd - be \geq 0;$$

$$ed - af \geq 0;$$

and

$$d \leq 0 .$$

Case 2. m is any positive integer

Using Cramer's Rule, the solution to the system $D\lambda^{o} = (P - A'\hat{\beta})$ is

$$\lambda_i^{o} = \left|\frac{M_i}{D}\right| \tag{2.41}$$

where $|M_i| = (-1)^{i+1}Q_1D^*(1,i) + (-1)^{i+2}Q_2D^*(2,i) + \ldots + (-1)^{i+m}Q_mD^*(m,i)$. This is the Laplace expansion of $\det D$ along the 1st column, where $(-1)^{i+j}D^*(j,i)$ is the classical adjoint of D, and Q_j is the jth component of the vector $(P - A'\hat{\beta})$.

When all constraints are violated, then $(P - A'\hat{\beta}) > 0$, so that a sufficient condition for each λ_i^{o} in (2.41) to be positive is

$$(-1)^{i+j}\, D^*(j,i) \geq 0 \text{ and strictly positive for some } j \text{ and } i .$$

We shall denote this condition as Condition A. Therefore whenever all constraints are violated and D satisfies Condition A, then $\lambda^{o} > 0$. From Lemma 2.6, then,

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}A\lambda^{o} \quad \text{where} \quad \lambda^{o} = \frac{1}{D}\begin{pmatrix} |M_1| \\ |M_2| \\ \vdots \\ |M_m| \end{pmatrix} \text{ using formula (2.41).}$$

Suppose now that m is any positive integer and that only some of the constraint equations are violated by $\hat{\beta}$. Let the constraint equations be partitioned into two sets as follows:

$$A'\beta \geq P \quad = \quad \begin{cases} A_1'\beta \geq P_1 \\ A_2'\beta \geq P_2 \end{cases}$$

A' is $m \times k$, A_1' is $m_1 \times k$, A_2' is $m_2 \times k$ and

$m_1 + m_2 = m$,

where

$$\begin{cases} A_1'\hat{\beta} \geq P_1 \\ A_2'\hat{\beta} < P_2 \end{cases} .$$

We must adjust $\hat{\beta}$ to $\tilde{\beta}$ so that $\tilde{\beta}$ satisfies the second set of constraints, while continuing to satisfy the first set of constraints.

Let

$$D = \left[\begin{array}{c|c} D_{11} \quad {}_{m_1 \times m_1} & D_{12} \quad {}_{m_1 \times m_2} \\ \hline D_{12}' \quad {}_{m_2 \times m_1} & D_{22} \quad {}_{m_2 \times m_2} \end{array}\right]_{m \times m}$$

represent the variance-covariance matrix of $(A_1'\hat{\beta}, A_2'\hat{\beta})$. Therefore we must solve the large system $D\lambda^o \geq P - A'\hat{\beta}$ subject to the cross-product equation $\lambda^{o'}(P - A'\tilde{\beta}) = 0$.

As in Case 3 when $m = 2$, it is impossible for $\lambda_2^o = 0$. Consequently, there are only two situations to consider.

<u>Subcase 1</u> $\lambda_1^o = 0$; $\lambda_2^o > 0$.

From (2.18) and (2.20), we must have

$$\begin{pmatrix} D_{11} & D_{12} \\ D_{12}' & D_{22} \end{pmatrix} \begin{pmatrix} 0 \\ \lambda_2^o \end{pmatrix} \begin{matrix} \geq \\ = \end{matrix} \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} ,$$

which can be written equivalently as

$$D_{12}\lambda_2^o \geq P_1 - A_1'\hat{\beta} \tag{2.42}$$

and

$$D_{22}\lambda_2^o = P_2 - A_2'\hat{\beta} \ . \tag{2.43}$$

If D_{22} satisfies Condition A, then from (2.43) $\lambda_2^o = D_{22}^{-1}(P_2 - A_2'\hat{\beta}) > 0$. Substituting for λ_2^o in (2.42),

$$D_{12}\, D_{22}^{-1}\,(P_2 - A_2'\hat{\beta}) \geq (P_1 - A_1'\hat{\beta}) \ . \tag{2.44}$$

Since $P_1 - A_1'\hat{\beta} \leq 0$ and $D_{22}^{-1}(P_2 - A_2'\hat{\beta}) > 0$, if all elements of D_{12} are nonnegative, then (2.44) is satisfied.

Consequently the closed form solution for β is

$$\begin{aligned} \tilde{\beta} &= \hat{\beta} + (X'X)^{-1}\,(A_1, A_2) \begin{pmatrix} 0 \\ D_{22}^{-1}\,(P_2 - A_2'\hat{\beta}) \end{pmatrix} \\ &= \hat{\beta} + (X'X)^{-1}\, A_2\, D_{22}^{-1}\,(P_2 - A_2'\hat{\beta}) \ . \end{aligned}$$

<u>Subcase 2</u> $\lambda_1^o > 0$; $\lambda_2^o > 0$.

From expressions (2.18) and (2.20), we must have

$$\begin{pmatrix} D_{11} & D_{12} \\ D_{12}' & D_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^o \\ \lambda_2^o \end{pmatrix} \begin{matrix} = \\ = \end{matrix} \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} ,$$

or

$$D_{11}\lambda_1^o + D_{12}\lambda_2^o = P_1 - A_1'\hat{\beta} \tag{2.45}$$

and

$$D_{12}'\lambda_1^o + D_{22}\lambda_2^o = P_2 - A_2'\hat{\beta} \; . \tag{2.46}$$

If all the elements of D_{11} are positive then $D_{11}\lambda_1^o > 0$. Since $\lambda_i^o > 0$ and $P_1 - A_1'\hat{\beta} \leq 0$, then (2.45) will not be satisfied when D_{12} consists of nonnegative elements.

Situations when (2.45) and (2.46) are satisfied can be found. However, the solutions for λ_1^o and λ_2^o depend on the structures of D_{11}, D_{12}, and D_{22}. Therefore, under Condition A the only possible general case is $\lambda_1^o = 0$ and $\lambda_2^o > 0$. We then have the following lemma:

<u>Lemma 2.7</u>. If D_{12}, the covariance matrix of $(A_1'\hat{\beta}, A_2'\hat{\beta})$ consists of all nonnegative elements and if D_{22} satisfies condition A, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 D_{22}^{-1} (P_2 - A_2'\hat{\beta}) \; .$$

Proof:

Since the formula for $\tilde{\beta}$ was derived from the Kuhn-Tucker conditions, we need only check that $\tilde{\beta}$ satisfies all constraints.

$$\begin{aligned} P_2 - A_2'\tilde{\beta} &= P_2 - A_2'\hat{\beta} - [A_2'(X'X)^{-1}A_2D_{22}^{-1}(P_2 - A_2'\hat{\beta})] \\ &= P_2 - A_2'\hat{\beta} - [A_2'(X'X)^{-1}A_2(A_2'(X'X)^{-1}A_2)^{-1}(P_2 - A_2'\hat{\beta})] \\ &= P_2 - A_2'\hat{\beta} - P_2 + A_2'\hat{\beta} = 0 \end{aligned}$$

$$P_1 - A_1'\tilde{\beta} = P_1 - A_1'\hat{\beta} - [A_1'(X'X)^{-1}A_2D_{22}^{-1}(P_2 - A_2'\hat{\beta})]$$

$$= P_1 - A_1'\hat{\beta} - [D_{12}D_{22}^{-1}(P_2 - A_2'\hat{\beta})]$$

$$= P_1 - A_1'\hat{\beta} - \text{(some positive quantity)} .$$

Since $P_1 - A_1'\hat{\beta}$ is negative, then $P_1 - A_1'\tilde{\beta}$ is negative, i.e., $A_1'\tilde{\beta} > P_1$. Consequently, all constraints are satisfied. $\square$

Numerical Example

Given the model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i ,$$

$i = 1, 2, 3,$ where $\varepsilon_i \sim NID(0, \sigma^2 I)$ and $\begin{pmatrix} 2\beta_0 + 6\beta_1 + 2\beta_2 \\ 2\beta_0 \qquad + 3\beta_2 \end{pmatrix} \geq \begin{pmatrix} 4 \\ 20 \end{pmatrix}$ and the data

X_1	1	2	3
X_2	1	4	9
Y	5	3	1

,

find the restricted least squares estimate for β .

For this data

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}; \qquad (X'X) = \begin{pmatrix} 3 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix};$$

$$(X'X)^{-1} = \begin{pmatrix} 19 & -21 & 5 \\ -21 & 24.5 & -6 \\ 5 & -6 & 1.5 \end{pmatrix}; \quad \text{and} \quad X'Y = \begin{pmatrix} 9 \\ 14 \\ 26 \end{pmatrix}.$$

Therefore

$$\hat{\beta} = (X'X)^{-1}X'Y = \begin{pmatrix} 19 & -21 & 5 \\ -21 & 24.5 & -6 \\ 5 & -6 & 1.5 \end{pmatrix}\begin{pmatrix} 9 \\ 14 \\ 26 \end{pmatrix} = \begin{pmatrix} 7 \\ -2 \\ 0 \end{pmatrix}.$$

Now

$$A' = \begin{pmatrix} 2 & 6 & 2 \\ 2 & 0 & 3 \end{pmatrix}; \qquad P = \begin{pmatrix} 4 \\ 20 \end{pmatrix};$$

$$A'\hat{\beta} = \begin{pmatrix} 2 & 6 & 2 \\ 2 & 0 & 3 \end{pmatrix}\begin{pmatrix} 7 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 14 \end{pmatrix} < \begin{pmatrix} 4 \\ 20 \end{pmatrix}.$$

Therefore both constraints are violated; this problem is an example of Case 2.

Next we must determine $\operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta})$.

$$D = A'(X'X)^{-1}A = \begin{pmatrix} 2 & 6 & 2 \\ 2 & 0 & 3 \end{pmatrix}\begin{pmatrix} 19 & -21 & 5 \\ -21 & 24.5 & -6 \\ 5 & -6 & 1.5 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 6 & 0 \\ 2 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 356 & -225 \\ -225 & 149.5 \end{pmatrix}.$$

Hence $d_{11} = 356$; $d_{12} = -225$; and $d_{22} = 149.5$. Since $\text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} < 0$, then from Table 2.2,

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} AD^{-1} (P - A'\hat{\beta})$$

$$= \begin{pmatrix} 7 \\ -2 \\ 0 \end{pmatrix} + \begin{pmatrix} 19 & -21 & 5 \\ -21 & 24.5 & -6 \\ 5 & -6 & 1.5 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 6 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} .0576 & .0866 \\ .0866 & .1371 \end{pmatrix} \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$

$$= \begin{pmatrix} 10.248 \\ -2.694 \\ -.166 \end{pmatrix}.$$

Notice that $\tilde{\beta}$ satisfies both constraints exactly, since

$$2\tilde{\beta}_0 + 6\tilde{\beta}_1 + 2\tilde{\beta}_2 = 4.0$$

and

$$2\tilde{\beta}_0 + 3\tilde{\beta}_2 = 20.0 .$$

CHAPTER III. CLOSED FORM SOLUTIONS FOR LEAST SQUARES ESTIMATORS WHEN THE PARAMETERS ARE SUBJECT TO LINEAR INTERVAL CONSTRAINTS

Introduction

In Chapter II we derived the least squares closed form solution for $\hat{\beta}$ when β is subject to constraints of the form $A'\beta \geq P$. A related but slightly more complicated problem is that of linear interval constraints on β, i.e., $P_0 \leq A'\beta \leq P_1$. For such constraints, whenever $\hat{\beta}$ is known, at most one of each pair of interval inequalities can be violated. In this chapter we will use this fact concerning the violated inequalities and the theory and results of Chapter II to derive the closed form solution for β, when β is restricted by linear interval constraints $P_0 \leq A'\beta \leq P_1$, where A' is either a $1 \times k$ or $2 \times k$ matrix. We conclude the chapter with two numerical examples.

A' is a $1 \times k$ Matrix

The constraints $P_0 \leq A'\beta \leq P_1$ can be written equivalently as

$$A'\beta \geq P_0$$

and $\hfill (3.1)$

$$-A'\beta \geq -P_1 \ .$$

For any value of $\hat{\beta}$, at least one of the inequalities of (3.1) is satisfied.

Case 1. $\hat{\beta}$ satisfies both restrictions

For $\lambda_1^o = 0$, $\lambda_2^o = 0$, and $\tilde{\beta} = \hat{\beta}$, the Kuhn-Tucker conditions (2.1)-(2.4) hold. Consequently the closed form solution for β is $\tilde{\beta} = \hat{\beta} = (X'X)^{-1}X'Y$.

Case 2. $\hat{\beta}$ violates the first inequality, i.e., $A'\hat{\beta} < P_0$

The only reasonable combination of λ_1^o and λ_2^o is $\lambda_1^o \geq 0$ and $\lambda_2^o = 0$. This corresponds to adjusting $\hat{\beta}$ to $\tilde{\beta}$ so that $\tilde{\beta}$ lies on the boundary of the constrained region, i.e., $A'\tilde{\beta} = P_0$. From (2.20), in order to satisfy the Kuhn-Tucker conditions, the following system must be solved:

$$\begin{pmatrix} A' \\ -A' \end{pmatrix} (X'X)^{-1} (A, -A) \begin{pmatrix} \lambda_1^o \\ 0 \end{pmatrix} \begin{matrix} = \\ \geq \end{matrix} \begin{pmatrix} P_0 - A'\hat{\beta} \\ -P_1 + A'\hat{\beta} \end{pmatrix} . \tag{3.2}$$

The system (3.2) implies that

$$A'(X'X)^{-1} A \lambda_1^o = P_0 - A'\hat{\beta} \quad \text{(positive)} \tag{3.3}$$

and

$$-A'(X'X)^{-1} A \lambda_1^o \geq -P_1 + A'\hat{\beta} \quad \text{(nonpositive).} \tag{3.4}$$

Solving for λ_1^o in (3.3), we have that

$$\lambda_1^o = \frac{P_0 - A'\hat{\beta}}{A'(X'X)^{-1}A} . \tag{3.5}$$

Substituting (3.5) for λ_1^o in (3.4), then

$$-A'(X'X)^{-1} A \left(\frac{P_0 - A'\hat{\beta}}{A'(X'X)^{-1}A} \right) = -P_0 + A'\hat{\beta} \geq -P_1 + A'\hat{\beta}$$

since, by assumption, $P_1 \geq P_0$. Therefore, (3.4) is always satisfied. The closed form solution for β from Table 2.1 is then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A \left(\frac{P_0 - A'\hat{\beta}}{A'(X'X)^{-1}A} \right) .$$

<u>Case 3.</u> <u>$\hat{\beta}$ violates the second inequality, i.e., $-A'\hat{\beta} < -P_1$</u>

Following the same derivation as in Case 2, the closed form solution for β is

$$\tilde{\beta} = \hat{\beta} - (X'X)^{-1} A \left(\frac{-P_1 + A'\hat{\beta}}{A'(X'X)^{-1}A} \right) .$$

Table 3.1 summarizes the results of this section.

Table 3.1. Constrained least squares solution for β where $P_0 \leq A'\beta \leq P_1$ and A' is a $1 \times k$ matrix

$A'\hat{\beta} < P_0$	$P_0 \leq A'\hat{\beta} \leq P_1$	$A'\hat{\beta} > P_1$
$\hat{\beta} + \dfrac{(X'X)^{-1} A(P_0 - A'\hat{\beta})}{A'(X'X)^{-1}A}$	$\hat{\beta}$	$\hat{\beta} - \dfrac{(X'X)^{-1} A(A'\hat{\beta} - P_1)}{A'(X'X)^{-1}A}$

A' is a 2 x k Matrix

For convenience, we can partition the matrices A and A' and vectors P_0 and P_1 as follows:

$$A = (A_1, A_2); \quad A' = \begin{pmatrix} A_1' \\ A_2' \end{pmatrix}; \quad P_0 = \begin{pmatrix} P_0 \\ P_2 \end{pmatrix}; \text{ and } P_1 = \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}.$$

The constraint inequalities can then be written in partitioned form as

$$P_0 \leq A_1'\beta \leq P_1 \tag{3.6}$$

and

$$P_2 \leq A_2'\beta \leq P_3 \quad . \tag{3.7}$$

Given the unrestricted least squares estimate $\hat{\beta}$, at most one inequality for $A_1'\hat{\beta}$ in (3.6) and one inequality for $A_2'\hat{\beta}$ in (3.7) can be violated. We will investigate the problem of simultaneous upper and lower bounds by considering the possible combinations of violated inequalities. We will adjust $\hat{\beta}$ to $\tilde{\beta}$ using the formulae developed in Chapter II so that the inequalities violated by $\hat{\beta}$ are satisfied by $\tilde{\beta}$, and then find conditions so that $\tilde{\beta}$ does not then violated the inequalities satisfied by $\hat{\beta}$.

Recall from Chapter II, $D = A'(X'X)^{-1}A$ is the variance-covariance matrix of $A_1'\hat{\beta}$ and $A_2'\hat{\beta}$. We again assume $(X'X)^{-1}$ exists and that A' is of full row rank, so that D is positive definite. Likewise, we define

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \quad \text{and} \quad D^{-1} = \frac{1}{d_{11}d_{22} - d_{12}^2}\begin{pmatrix} d_{22} & -d_{12} \\ -d_{12} & d_{11} \end{pmatrix}.$$

Det $D = d_{11}d_{22} - d_{12}^2 > 0$ since D is positive definite.

<u>Case 1.</u> <u>$\hat{\beta}$ satisfies all restrictions, i.e., $P_0 \leq A_1'\hat{\beta} \leq P_1$ and $P_2 \leq A_2'\hat{\beta} \leq P_3$</u>

Rewriting the interval inequalities as four single inequalities, i.e., $A_1'\beta \geq P_0$; $-A_1'\beta \geq -P_1$; $A_2'\beta \geq P_2$; and $-A_2'\beta \geq -P_3$, then the vector $\lambda^o = 0$ and $\hat{\beta}$ clearly satisfy the Kuhn-Tucker conditions. Hence, the closed form solution for β is $\tilde{\beta} = \hat{\beta} = (X'X)^{-1}X'Y$.

<u>Case 2.</u> <u>$\hat{\beta}$ violates both upper bounds, i.e., $A_1'\hat{\beta} > P_1$ and $A_2'\hat{\beta} > P_3$</u>

In this situation, we need to adjust $\hat{\beta}$ to $\tilde{\beta}$ so that

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} P_1 - \alpha_1(P_1 - P_0) \\ P_3 - \alpha_2(P_3 - P_2) \end{pmatrix} , \tag{3.8}$$

where α_1 and $\alpha_2 \in [0, 1]$.

In order to appeal to Tables 2.2, 2.3, and 2.4, we must rewrite this particular case in the framework of Chapter II. We will, therefore, temporarily ignore the lower inequalities of each interval constraint, so that, as in Chapter II, we have only two inequality constraints. Moreover, we must make sure that the $\tilde{\beta}$ from Tables 2.2, 2.3, or 2.4 continues to satisfy equation (3.8).

Now, rewriting this case in the framework of Chapter II:

$$\left.\begin{aligned} -A_1'\beta &\geq -P_1 \\ -A_2'\beta &\geq -P_3 \end{aligned}\right\} \Longleftrightarrow \left\{\begin{aligned} \bar{A}_1'\beta &\geq -P_1 \\ \bar{A}_2'\beta &\geq -P_3 \end{aligned}\right.$$

where both constraints are violated by $\hat{\beta}$.

In order to use Table 2.2, we need to find the elements of

$$\bar{D} = \begin{pmatrix} \bar{A}_1' \\ \bar{A}_2' \end{pmatrix} (X'X)^{-1} (\bar{A}_1, \bar{A}_2)$$

in terms of the elements of

$$D = \begin{pmatrix} A_1' \\ A_2' \end{pmatrix} (X'X)^{-1} (A_1, A_2) .$$

Here

$$\bar{D} = \begin{pmatrix} \overline{d_{11}} & \overline{d_{12}} \\ \overline{d_{12}} & \overline{d_{22}} \end{pmatrix}$$

$$= \begin{pmatrix} -A_1' \\ -A_2' \end{pmatrix} (X'X)^{-1} (-A_1, -A_2)$$

$$= \begin{pmatrix} A_1' \\ A_2' \end{pmatrix} (X'X)^{-1} (A_1, A_2)$$

$$= D$$

$$= \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} .$$

a) If $\operatorname{cov}(\bar{A}_1'\hat{\beta}, \bar{A}_2'\hat{\beta}) = \operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} \leq 0$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}(\bar{A})\bar{D}^{-1} \begin{pmatrix} -P_1 - \bar{A}_1'\hat{\beta} \\ -P_3 - \bar{A}_2'\hat{\beta} \end{pmatrix}$$

$$= \hat{\beta} - (X'X)^{-1}AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

from Table 2.2.

Therefore

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = A'\tilde{\beta}$$

$$= A'\hat{\beta} - A'(X'X)^{-1}AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

$$= A'\hat{\beta} - \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

$$= \begin{pmatrix} P_1 \\ P_3 \end{pmatrix} . \tag{3.9}$$

Putting equations (3.8) and (3.9) together, we have that

$$\begin{pmatrix} P_1 - \alpha_1(P_1 - P_0) \\ P_3 - \alpha_2(P_3 - P_2) \end{pmatrix} = \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}, \tag{3.10}$$

where α_1 and $\alpha_2 \varepsilon [0, 1]$.

Hence (3.10) implies that $\alpha_1 = \alpha_2 = 0$ and $A'\tilde{\beta} = \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$, where

$\tilde{\beta} = \hat{\beta} - (X'X)^{-1} AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$.

b) If $\operatorname{cov}(\bar{A}_1'\hat{\beta}, \bar{A}_2'\hat{\beta}) = \operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} > 0$ and

$\frac{d_{12}}{d_{22}} < \frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} < \frac{d_{11}}{d_{12}}$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} \bar{A}\, \bar{D}^{-1} \begin{pmatrix} -P_1 - \bar{A}_1'\hat{\beta} \\ -P_3 - \bar{A}_2'\hat{\beta} \end{pmatrix}$$

$$= \hat{\beta} - (X'X)^{-1} AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

from Table 2.2.

Therefore

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = A'\tilde{\beta}$$

$$= A'\hat{\beta} - A'(X'X)^{-1}AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

$$= A'\hat{\beta} - \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

$$= \begin{pmatrix} P_1 \\ P_3 \end{pmatrix} . \tag{3.11}$$

Hence, as above, equations (3.11) and (3.8) imply that $\alpha_1 = \alpha_2 = 0$ and $A'\tilde{\beta} = \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$, where $\tilde{\beta} = \hat{\beta} - (X'X)^{-1}AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$.

c) If $\text{cov}(\bar{A}_1'\hat{\beta}, \bar{A}_2'\hat{\beta}) = \text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} > 0$ and $\dfrac{d_{12}}{d_{22}} \geq \dfrac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} \bar{A}_2 \left(\frac{-P_3 - \bar{A}_2'\hat{\beta}}{\bar{d}_{22}} \right)$$

$$= \hat{\beta} - (X'X)^{-1} A_2 \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}} \right)$$

from Table 2.2.

Therefore

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} A_1'\hat{\beta} - \dfrac{A_1'(X'X)^{-1}A_2(-P_3 + A_2'\hat{\beta})}{d_{22}} \\ A_2'\hat{\beta} - \dfrac{A_2'(X'X)^{-1}A_2(-P_3 + A_2'\hat{\beta})}{d_{22}} \end{pmatrix}$$

$$= \begin{pmatrix} A_1'\hat{\beta} - \dfrac{d_{12}}{d_{22}} (-P_3 + A_2'\hat{\beta}) \\ \\ P_3 \end{pmatrix} . \qquad (3.12)$$

In this case, when (3.12) and (3.8) are set equal,

$$\begin{pmatrix} P_1 - \alpha_1(P_1 - P_0) \\ \\ P_3 - \alpha_2(P_3 - P_2) \end{pmatrix} = \begin{pmatrix} A_1'\hat{\beta} - \dfrac{d_{12}}{d_{22}} (-P_3 + A_2'\hat{\beta}) \\ \\ P_3 \end{pmatrix} \qquad \begin{matrix} (3.13) \\ \\ (3.14) \end{matrix}$$

where α_1 and $\alpha_2 \varepsilon [0, 1]$.

From (3.14), $\alpha_2 = 0$.

Rewriting (3.13),

$$\alpha_1(P_1 - P_0) = \frac{d_{12}}{d_{22}} (-P_3 + A_2'\hat{\beta}) - (-P_1 + A_1'\hat{\beta}) . \qquad (3.15)$$

Before continuing, we recall that $(P_1 - P_0)$, d_{12}, d_{22}, $(-P_3 + A_2'\hat{\beta})$, and $(-P_1 + A_1'\hat{\beta})$ are positive.

Since $\alpha_1 \geq 0$, then from (3.15)

$$\frac{d_{12}}{d_{22}} (-P_3 + A_2'\hat{\beta}) \geq (-P_1 + A_1'\hat{\beta}) ,$$

or

$$\frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \leq \frac{d_{12}}{d_{22}} , \qquad (3.16)$$

which is the condition assumed from Table 2.2.

Also, since $\alpha_1 \leq 1$, then from (3.15),

$$P_1 - P_0 \geq \frac{d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta}) - (-P_1 + A_1'\hat{\beta})$$

or

$$P_1 - P_0 + (-P_1 + A_1'\hat{\beta}) \geq \frac{d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta})$$

or

$$\frac{-P_0 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \geq \frac{d_{12}}{d_{22}} . \tag{3.17}$$

Consequently, whenever $d_{12} > 0$ and

$$\frac{-P_0 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \geq \frac{d_{12}}{d_{22}} \geq \frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$$

then

$$\tilde{\beta} = \hat{\beta} - (X'X)^{-1} A_2 \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}} \right).$$

d) If $\text{cov}(\bar{A}_1'\hat{\beta}, \bar{A}_2'\hat{\beta}) = \text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} > 0$ and $\frac{d_{11}}{d_{12}} \leq \frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} \bar{A}_1 \left(\frac{-P_1 - \bar{A}_1'\hat{\beta}}{\bar{d}_{11}} \right)$$

$$= \hat{\beta} - (X'X)^{-1} A_1 \left(\frac{-P_1 + A_1'\hat{\beta}}{d_{11}} \right)$$

from Table 2.2.

Therefore

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} A_1'\hat{\beta} - \dfrac{A_1'(X'X)^{-1} A_1(-P_1 + A_1'\hat{\beta})}{d_{11}} \\ A_2'\hat{\beta} - \dfrac{A_2'(X'X)^{-1} A_1(-P_1 + A_1'\hat{\beta})}{d_{11}} \end{pmatrix}$$

$$= \begin{pmatrix} P_1 \\ A_2'\hat{\beta} - \dfrac{d_{12}}{d_{11}} (-P_1 + A_1'\hat{\beta}) \end{pmatrix}. \tag{3.18}$$

Similarly, when (3.18) and (3.8) are set equal,

$$\begin{pmatrix} P_1 - \alpha_1(P_1 - P_0) \\ P_3 - \alpha_2(P_3 - P_2) \end{pmatrix} = \begin{pmatrix} P_1 \\ A_2'\hat{\beta} - \dfrac{d_{12}}{d_{11}} (-P_1 + A_1'\hat{\beta}) \end{pmatrix} \qquad \begin{matrix} (3.19) \\ (3.20) \end{matrix}$$

where α_1 and $\alpha_2 \in [0, 1]$.

From (3.19), $\alpha_1 = 0$.

Rewriting (3.20),

$$\alpha_2(P_3 - P_2) = \frac{d_{12}}{d_{11}}(-P_1 + A_1'\hat{\beta}) - (-P_3 + A_2'\hat{\beta}) . \qquad (3.21)$$

Now $(P_3 - P_2)$, d_{12}, d_{11}, $(-P_1 + A_1'\hat{\beta})$, and $(-P_2 + A_2'\hat{\beta}) > 0$. Since $\alpha_2 \geq 0$, then from (3.21)

$$\frac{d_{12}}{d_{11}}(-P_1 + A_1'\hat{\beta}) \geq (-P_3 + A_2'\hat{\beta})$$

or

$$\frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \geq \frac{d_{11}}{d_{12}} , \qquad (3.22)$$

which is the assumed condition given in Table 2.2.

Also, since $\alpha_2 \leq 1$, then from (3.21),

$$P_3 - P_2 \geq \frac{d_{12}}{d_{11}}(-P_1 + A_1'\hat{\beta}) - (-P_3 + A_2'\hat{\beta})$$

or

$$P_3 - P_2 + (-P_3 + A_2'\hat{\beta}) \geq \frac{d_{12}}{d_{11}}(-P_1 + A_1'\hat{\beta})$$

or

$$\frac{-P_1 + A_1'\hat{\beta}}{-P_2 + A_2'\hat{\beta}} \leq \frac{d_{11}}{d_{12}} \quad \text{since} \quad -P_2 + A_2'\hat{\beta} \text{ is positive.} \qquad (3.23)$$

Therefore, if $d_{12} > 0$ and

$$\frac{-P_1 + A_1'\hat{\beta}}{-P_2 + A_2'\hat{\beta}} \leq \frac{d_{11}}{d_{12}} \leq \frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$$

then

$$\tilde{\beta} = \hat{\beta} - (X'X)^{-1} A_1 \left(\frac{-P_1 + A_1'\hat{\beta}}{d_{11}} \right) .$$

Tables 3.2 and 3.3 summarize the situations when either both upper or both lower bounds are violated by $\hat{\beta}$.

Case 3. $\hat{\beta}$ violates the first lower and second upper bounds, i.e., $A_1'\hat{\beta} < P_0$ and $A_2'\hat{\beta} > P_3$

We need to adjust $\hat{\beta}$ to $\tilde{\beta}$ so that

$$\begin{pmatrix} A_1'\tilde{\beta} \\ -A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} P_1 + \alpha_1(P_1 - P_0) \\ -P_3 + \alpha_2(P_3 - P_2) \end{pmatrix} \tag{3.24}$$

where α_1 and $\alpha_2 \in [0, 1]$.

Rewriting this case in the light of Chapter II:

$$\left.\begin{matrix} A_1'\beta \geq P_0 \\ -A_2'\beta \geq -P_3 \end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix} \bar{A}_1'\beta \geq P_0 \\ \bar{A}_2'\beta \geq -P_3 \end{matrix}\right.$$

where both constraints are violated by $\hat{\beta}$. Now

$$\bar{D} = \bar{A}' (X'X)^{-1} \bar{A} = \begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} (X'X)^{-1} (A_1, -A_2)$$

$$= \begin{pmatrix} d_{11} & -d_{12} \\ -d_{12} & d_{22} \end{pmatrix} .$$

Also

$$\bar{D}^{-1} = \frac{1}{d_{11}d_{22} - d_{12}^2} \begin{pmatrix} d_{22} & d_{12} \\ d_{12} & d_{11} \end{pmatrix} .$$

a) If $\operatorname{cov}(\bar{A}_1'\hat{\beta}, \bar{A}_2'\hat{\beta}) = \operatorname{cov}(A_1'\hat{\beta}, -A_2'\hat{\beta}) = -d_{12} \leq 0$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} (A_1, -A_2)\bar{D}^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

from Table 2.2.

Therefore

$$\begin{pmatrix} A_1'\tilde{\beta} \\ -A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} \tilde{\beta}$$

$$= \begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} \hat{\beta} + \begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} (X'X)^{-1}(A_1, -A_2)\bar{D}^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

$$= \begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} \hat{\beta} + \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$$

$$= \begin{pmatrix} P_0 \\ -P_3 \end{pmatrix} . \tag{3.25}$$

Along with (3.24), (3.25) implies that $\alpha_1 = \alpha_2 = 0$. Thus

$$\begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} \tilde{\beta} = \begin{pmatrix} P_0 \\ -P_3 \end{pmatrix}$$

where

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}(A_1, -A_2)\, \bar{D}^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix} .$$

b) If $\operatorname{cov}(\bar{A}_1'\hat{\beta}, -\bar{A}_2'\hat{\beta}) = \operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = -d_{12} > 0$ and

$$\frac{-d_{12}}{d_{22}} < \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} < \frac{d_{11}}{-d_{12}} ,$$

then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} (A_1, -A_2)\, \bar{D}^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix} ,$$

from Table 2.2.

This is identical to the above situation, so $\alpha_1 = \alpha_2 = 0$ and

$$\begin{pmatrix} A_1' \\ -A_2' \end{pmatrix} \tilde{\beta} = \begin{pmatrix} P_0 \\ -P_3 \end{pmatrix} .$$

c) If $\operatorname{cov}(\overline{A}_1'\hat{\beta}, -\overline{A}_2'\hat{\beta}) = \operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = -d_{12} > 0$ and

$$\frac{-d_{12}}{d_{22}} \geq \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \text{ , then}$$

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} (-A_2) \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}} \right)$$

from Table 2.2.

Therefore

$$A_1'\tilde{\beta} = A_1'\hat{\beta} - A_1'(X'X)^{-1} (A_2) \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}} \right)$$

$$= A_1'\hat{\beta} + \frac{-d_{12}}{d_{22}} (-P_3 + A_2'\hat{\beta}) \qquad (3.26)$$

and

$$-A_2'\tilde{\beta} = -A_2'\hat{\beta} - A_2'(X'X)^{-1}(-A_2) \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}} \right)$$

$$= -A_2'\hat{\beta} + (-P_3 + A_2'\hat{\beta}) = -P_3 \quad . \qquad (3.27)$$

Along with (3.24), (3.27) implies that $\alpha_2 = 0$.

Putting (3.26) together with (3.24),

$$A_1'\hat{\beta} + \frac{-d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta}) = P_0 + \alpha_1(P_1 - P_0) ,$$

which can be rewritten as

$$\alpha_1(P_1 - P_0) = -(P_0 - A_1'\hat{\beta}) + \frac{-d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta}) . \qquad (3.28)$$

Now $-d_{12}$, d_{22}, $P_1 - P_0$, $(P_0 - A_1'\hat{\beta})$ and $(-P_3 + A_2'\hat{\beta})$ are positive, and since $\alpha_1 \geq 0$, then (3.28) implies that

$$\frac{-d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta}) \geq (P_0 - A_1'\hat{\beta})$$

or

$$\frac{-d_{12}}{d_{22}} \geq \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \qquad (3.29)$$

which is the condition assumed from Table 2.2.

Also $\alpha_1 \leq 1$, so (3.28) implies that

$$(P_1 - P_0) \geq -(P_0 - A_1'\hat{\beta}) + \frac{-d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta})$$

or

$$P_1 - P_0 + (P_0 - A_1'\hat{\beta}) \geq \frac{-d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta})$$

or

$$P_1 - A_1'\hat{\beta} \geq \frac{-d_{12}}{d_{22}}(-P_3 + A_2'\hat{\beta})$$

or

$$\frac{P_1 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \geq \frac{-d_{12}}{d_{22}} \quad . \tag{3.30}$$

Consequently, whenever

$$\frac{P_1 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \geq \frac{-d_{12}}{d_{22}} \geq \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$$

then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}(-A_2)\left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}}\right)$$

$$= \hat{\beta} - (X'X)^{-1} A_2 \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}}\right).$$

d) If $\operatorname{cov}(\bar{A}_1'\hat{\beta}, -\bar{A}_2'\hat{\beta}) = \operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = -d_{12} > 0$, and

$$\frac{d_{11}}{-d_{22}} \leq \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} ,$$

then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}(A_1)\left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}}\right)$$

from Table 2.2.

Therefore

$$\begin{aligned} A_1'\tilde{\beta} &= A_1'\hat{\beta} + A_1'(X'X)^{-1}(A_1)\left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}}\right) \\ &= A_1'\hat{\beta} + P_0 - A_1'\hat{\beta} \\ &= P_0 \end{aligned} \tag{3.31}$$

and

$$\begin{aligned} -A_2'\tilde{\beta} &= -A_2'\hat{\beta} - A_2'(X'X)^{-1}(A_1)\left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}}\right) \\ &= -A_2'\hat{\beta} + \frac{-d_{12}}{d_{11}}(P_0 - A_1'\hat{\beta}) \ . \end{aligned} \tag{3.22}$$

Together, (3.31) and (3.24) imply that $\alpha_1 = 0$.

Expressions (3.32) and (3.24) imply that

$$-A_2'\hat{\beta} + \frac{-d_{12}}{d_{11}}(P_0 - A_1'\hat{\beta}) = -P_3 + \alpha_2(P_3 - P_2) \quad . \tag{3.33}$$

Recall that $-d_{12}$, d_{22}, $P_3 - P_2$, $(P_0 - A_1'\hat{\beta})$ and $(-P_2 + A_2'\hat{\beta})$ are positive, so since $\alpha_2 \geq 0$, then from (3.33)

$$\frac{-d_{12}}{d_{11}}(P_0 - A_1'\hat{\beta}) \geq -P_3 + A_2'\hat{\beta}$$

or

$$\frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \geq \frac{d_{11}}{-d_{12}} \tag{3.34}$$

which is the condition assumed from Table 2.2. Also, since $\alpha_2 \leq 1$, then (3.33) implies that

$$(P_3 - P_2) \geq P_3 - A_2'\hat{\beta} + \frac{-d_{12}}{d_{11}}(P_0 - A_1'\hat{\beta})$$

or

$$P_3 - P_2 - P_3 + A_2'\hat{\beta} \geq \frac{-d_{12}}{d_{11}}(P_0 - A_1'\hat{\beta})$$

or

$$A_2'\hat{\beta} - P_2 \geq \frac{-d_{12}}{d_{11}}(P_0 - A_1'\hat{\beta})$$

or

$$\frac{d_{11}}{-d_{12}} \geq \frac{P_0 - A_1'\hat{\beta}}{-P_2 + A_2'\hat{\beta}} \quad \text{since } -P_2 + A_2'\hat{\beta} \text{ is positive.} \tag{3.35}$$

Consequently, whenever

$$\frac{P_0 - A_1'\hat{\beta}}{-P_2 + A_2'\hat{\beta}} \leq \frac{d_{11}}{-d_{12}} \leq \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$$

then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_1 \left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}} \right) .$$

Tables 3.4 and 3.5 summarize the situations when one upper and one lower bound are violated by $\hat{\beta}$.

Case 4. $\hat{\beta}$ violates only the second lower bound, i.e. $A_2'\hat{\beta} < P_2$

In order to use the results of Chapter II, we will first assume we have only the restrictions

$$A_1'\beta \geq P_0$$

and

$$A_2'\beta \geq P_2$$

where the second constraint is violated by $\hat{\beta}$. We will need to adjust $\hat{\beta}$ to $\tilde{\beta}$ so that

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} P_0 + \alpha_1(P_1 - P_0) \\ P_2 + \alpha_2(P_3 - P_2) \end{pmatrix} \qquad (3.36)$$

where α_1 and $\alpha_2 \in [0, 1]$.

a) If $\text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} \geq 0$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

from Table 2.4.

Therefore

$$A_1'\tilde{\beta} = A_1'\hat{\beta} + A_1'(X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

$$= A_1'\hat{\beta} + \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta}) \tag{3.37}$$

and

$$A_2'\tilde{\beta} = A_2'\hat{\beta} + A_2'(X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

$$= A_2'\hat{\beta} + (P_2 - A_2'\hat{\beta}) = P_2 \ . \tag{3.38}$$

Equations (3.38) and (3.36) together imply that $\alpha_2 = 0$.
Expressions (3.37) and (3.36) imply that

$$A_1'\hat{\beta} + \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta}) = P_0 + \alpha_1(P_1 - P_0)$$

which can be written equivalently as

$$-P_0 + A_1'\hat{\beta} + \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta}) = \alpha_1(P_1 - P_0) \ . \tag{3.39}$$

Notice that d_{12}, d_{22}, $(P_2 - A_2'\hat{\beta})$, $P_1 - P_0$, and $(-P_0 + A_1'\hat{\beta})$ are positive, so that knowing $\alpha_1 \geq 0$ is of little value in this situation. However, since $\alpha_1 \leq 1$, then from (3.39)

$$-P_0 + A_1'\hat{\beta} + \frac{d_{12}}{d_{22}}(P_2 - A_2'\hat{\beta}) \leq (P_1 - P_0)$$

or

$$\frac{d_{12}}{d_{22}}(P_2 - A_2'\hat{\beta}) \leq P_1 - A_1'\hat{\beta}$$

or

$$\frac{d_{12}}{d_{22}} \leq \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} \quad . \tag{3.40}$$

Hence, if

$$\frac{d_{12}}{d_{22}} \leq \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$$

then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right) .$$

b) If $\operatorname{cov}(A_1'\hat{\beta}, A_2'\hat{\beta}) = d_{12} < 0$ and $\dfrac{d_{12}}{d_{22}} \geq \dfrac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

from Table 2.4.

Consequently

$$A_1'\tilde{\beta} = A_1'\hat{\beta} + A_1'(X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

$$= A_1'\hat{\beta} + \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta}) \tag{3.41}$$

and

$$A_2'\tilde{\beta} = A_2'\hat{\beta} + A_2'(X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

$$= A_2'\hat{\beta} + (P_2 - A_2'\hat{\beta}) = P_2 \quad . \tag{3.42}$$

Once again (3.42) and (3.36) imply $\alpha_2 = 0$. Putting (3.41) together with (3.36) ,

$$A_1'\hat{\beta} + \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta}) = P_0 + \alpha_1(P_1 - P_0) \ . \tag{3.43}$$

Now since $(-P_0 + A_1'\hat{\beta})$, d_{22}, $(P_2 - A_2'\hat{\beta})$ and $P_1 - P_0$ are positive, then since $\alpha_1 \geq 0$ equation (3.43) implies that

$$-P_0 + A_1'\hat{\beta} \geq \frac{|d_{12}|}{d_{22}} (P_2 - A_2'\hat{\beta})$$

or

$$\frac{-P_0 + A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} \geq \frac{|d_{12}|}{d_{22}}$$

or

$$\frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} \leq \frac{d_{12}}{d_{22}} , \qquad (3.44)$$

which is the condition assumed from Table 2.4.

Also $\alpha_1 \leq 1$, so (3.43) implies that

$$- P_0 + A_1'\hat{\beta} + \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta}) \leq P_1 - P_0$$

or

$$P_1 - A_1'\hat{\beta} \geq \frac{d_{12}}{d_{22}} (P_2 - A_2'\hat{\beta})$$

or

$$\frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} \geq \frac{d_{12}}{d_{22}} , \qquad (3.45)$$

which is always true since d_{12} is negative and $(P_1 - A_1'\hat{\beta})$, $(P_2 - A_2'\hat{\beta})$, and d_{22} are positive. Hence, if

$$\frac{d_{12}}{d_{22}} \geq \frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$$

whenever $d_{12} < 0$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right) .$$

c) If $\operatorname{cov}(A_1'\hat{\beta},\ A_2'\hat{\beta}) = d_{12} < 0$ and

$$\frac{d_{12}}{d_{22}} < \frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$$

then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}\, AD^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$$

from Table 2.4.

Consequently

$$\begin{aligned} \begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} &= A'\tilde{\beta} \\ &= A'\hat{\beta} + A'(X'X)^{-1}\, AD^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} \\ &= \begin{pmatrix} P_0 \\ P_2 \end{pmatrix} . \end{aligned} \tag{3.46}$$

Together with (3.36), expression (3.46) implies that $\alpha_1 = \alpha_2 = 0$.

Thus

$$\begin{pmatrix} A_1'\tilde{\beta} \\ A_2'\tilde{\beta} \end{pmatrix} = \begin{pmatrix} P_0 \\ P_2 \end{pmatrix},$$

where

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}AD^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix} .$$

Tables 3.6-3.9 summarize the four situations when only one inequality is violated by $\hat{\beta}$.

In Tables 3.2-3.9,

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} = A'(X'X)^{-1}A$$

and

$$\bar{D} = \begin{pmatrix} d_{11} & -d_{12} \\ -d_{12} & d_{22} \end{pmatrix} .$$

Table 3.2. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ and $A'\hat{\beta} > \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$

Condition	$d_{12} \leq 0$	$d_{12} > 0$
none	$\hat{\beta} - (X'X)^{-1}AD^{-1}\begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$	
$\frac{d_{12}}{d_{22}} < \frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} < \frac{d_{11}}{d_{12}}$		$\hat{\beta} - (X'X)^{-1} AD^{-1} \begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$
$\frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \leq \frac{d_{12}}{d_{22}} \leq \frac{-P_0 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$		$\hat{\beta} - (X'X)^{-1} A_2 \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}}\right)$
$\frac{-P_1 + A_1'\hat{\beta}}{-P_2 + A_2'\hat{\beta}} \leq \frac{d_{11}}{d_{12}} \leq \frac{-P_1 + A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$		$\hat{\beta} - (X'X)^{-1} A_1 \left(\frac{-P_1 + A_1'\hat{\beta}}{d_{11}}\right)$

Table 3.3. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ and $A'\hat{\beta} < \begin{pmatrix} P_0 \\ P_2 \end{pmatrix}$

Condition	$d_{12} \leq 0$	$d_{12} > 0$
none	$\hat{\beta} + (X'X)^{-1}AD^{-1}\begin{pmatrix} P_0 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$	
$\frac{d_{12}}{d_{22}} < \frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} < \frac{d_{11}}{d_{12}}$		$\hat{\beta} + (X'X)^{-1}AD^{-1}\begin{pmatrix} P_0 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$
$\frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} \leq \frac{d_{12}}{d_{22}} \leq \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$		$\hat{\beta} + (X'X)^{-1}A_2\left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}}\right)$
$\frac{P_0 - A_1'\hat{\beta}}{P_3 - A_2'\hat{\beta}} \leq \frac{d_{11}}{d_{12}} \leq \frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$		$\hat{\beta} + (X'X)^{-1}A_1\left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}}\right)$

Table 3.4. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$, $A_1'\hat{\beta} < P_0$ and $A_2'\hat{\beta} > P_3$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
none		$\hat{\beta} + (X'X)^{-1}(A_1, -A_2)\bar{D}^{-1}\begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$
$\frac{-d_{12}}{d_{22}} < \frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} < \frac{d_{11}}{-d_{12}}$	$\hat{\beta} + (X'X)^{-1}(A_1, -A_2)\bar{D}^{-1}\begin{pmatrix} P_0 - A_1'\hat{\beta} \\ -P_3 + A_2'\hat{\beta} \end{pmatrix}$	
$\frac{P_0 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}} \leq \frac{-d_{12}}{d_{22}} \leq \frac{P_1 - A_1'\hat{\beta}}{-P_3 + A_2'\hat{\beta}}$	$\hat{\beta} - (X'X)^{-1} A_2 \left(\frac{-P_3 + A_2'\hat{\beta}}{d_{22}} \right)$	
$\frac{-P_3 + A_2'\hat{\beta}}{P_0 - A_1'\hat{\beta}} \leq \frac{-d_{12}}{d_{11}} \leq \frac{-P_2 + A_2'\hat{\beta}}{P_0 - A_1'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} A_1 \left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}} \right)$	

Table 3.5. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$, $A_1'\hat{\beta} > P_1$ and $A_2'\hat{\beta} < P_2$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
none		$\hat{\beta} + (X'X)^{-1}(-A_1, A_2)\overline{D}^{-1}\begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$
$\frac{-d_{12}}{d_{22}} < \frac{-P_1 + A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} < \frac{d_{11}}{-d_{12}}$	$\hat{\beta} + (X'X)^{-1}(-A_1, A_2)\overline{D}^{-1}\begin{pmatrix} -P_1 + A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$	
$\frac{P_2 - A_2'\hat{\beta}}{-P_1 + A_1'\hat{\beta}} \leq \frac{-d_{12}}{d_{11}} \geq \frac{P_3 - A_2'\hat{\beta}}{-P_1 + A_1'\hat{\beta}}$	$\hat{\beta} - (X'X)^{-1}A_1\left(\frac{-P_1 + A_1'\hat{\beta}}{d_{11}}\right)$	
$\frac{-P_1 + A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} \leq \frac{-d_{12}}{d_{22}} \leq \frac{-P_0 + A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1}A_2\left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}}\right)$	

Table 3.6. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ and $A_1'\hat{\beta} < P_0$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
$\frac{d_{12}}{d_{11}} \leq \frac{P_3 - A_2'\hat{\beta}}{P_0 - A_1'\hat{\beta}}$		$\hat{\beta} + (X'X)^{-1} A_1 \left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}} \right)$
$\frac{d_{12}}{d_{11}} \geq \frac{P_2 - A_2'\hat{\beta}}{P_0 - A_1'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} A_1 \left(\frac{P_0 - A_1'\hat{\beta}}{d_{11}} \right)$	
$\frac{d_{12}}{d_{11}} < \frac{P_2 - A_2'\hat{\beta}}{P_0 - A_1'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} A D^{-1} \begin{pmatrix} P_0 - A_1'\beta \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$	

Table 3.7. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ and $A_1'\hat{\beta} > P_1$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
$\dfrac{d_{12}}{d_{11}} \leq \dfrac{P_2 - A_2'\hat{\beta}}{P_1 - A_1'\hat{\beta}}$		$\hat{\beta} + (X'X)^{-1} A_1 \left(\dfrac{P_1 - A_1'\hat{\beta}}{d_{11}} \right)$
$\dfrac{d_{12}}{d_{11}} \geq \dfrac{P_3 - A_2'\hat{\beta}}{P_1 - A_1'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} A_1 \left(\dfrac{P_1 - A_1'\hat{\beta}}{d_{11}} \right)$	
$\dfrac{d_{12}}{d_{11}} < \dfrac{P_3 - A_2'\hat{\beta}}{P_1 - A_1'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} AD^{-1} \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_3 - A_2'\hat{\beta} \end{pmatrix}$	

Table 3.8. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ and $A_2'\hat{\beta} < P_2$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
$\frac{d_{12}}{d_{22}} \leq \frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$		$\hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$
$\frac{d_{12}}{d_{22}} \geq \frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$	
$\frac{d_{12}}{d_{22}} < \frac{P_0 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} AD^{-1} \begin{pmatrix} P_0 - A_1'\hat{\beta} \\ P_2 - A_2'\hat{\beta} \end{pmatrix}$	

Table 3.9. Constrained least squares solution for β where $\begin{pmatrix} P_0 \\ P_2 \end{pmatrix} \leq A'\beta \leq \begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ and $A_2'\hat{\beta} > P_3$

Condition	$d_{12} < 0$	$d_{12} \geq 0$
$\dfrac{d_{12}}{d_{22}} \leq \dfrac{P_0 - A_1'\hat{\beta}}{P_3 - A_2'\hat{\beta}}$		$\hat{\beta} + (X'X)^{-1} A_2 \left(\dfrac{P_3 - A_2'\hat{\beta}}{d_{22}} \right)$
$\dfrac{d_{12}}{d_{22}} \geq \dfrac{P_1 - A_1'\hat{\beta}}{P_3 - A_2'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} A_2 \left(\dfrac{P_3 - A_2'\hat{\beta}}{d_{22}} \right)$	
$\dfrac{d_{12}}{d_{22}} < \dfrac{P_1 - A_1'\hat{\beta}}{P_3 - A_2'\hat{\beta}}$	$\hat{\beta} + (X'X)^{-1} AD^{-1} \begin{pmatrix} P_1 - A_1'\hat{\beta} \\ P_3 - A_2'\hat{\beta} \end{pmatrix}$	

Numerical Examples

Example 1

Given the model $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i = 1, 2, 3$, where $\epsilon_i \sim NID(0, \sigma^2 I)$, and

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix} \leq \begin{pmatrix} \beta_0 + 2\beta_1 \\ 2\beta_0 + \beta_1 \end{pmatrix} \leq \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

and the data

X	1	2	3
Y	1	3	2

find the restricted least squares estimate for β .

For this data

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}; \qquad (X'X) = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix};$$

$$(X'X)^{-1} = \begin{pmatrix} 7/3 & -1 \\ -1 & 1/2 \end{pmatrix}; \quad \text{and} \quad X'Y = \begin{pmatrix} 6 \\ 13 \end{pmatrix}.$$

Consequently

$$\hat{\beta} = (X'X)^{-1} X'Y = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix}.$$

Now

$$A_1'\hat{\beta} = (1 \quad 2)\begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = 2$$

and

$$A_2'\hat{\beta} = (2 \quad 1)\begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = 5/2 .$$

Therefore, $A_2'\hat{\beta}$ violates the second lower bound, and we can use Table 3.6.

Next we must determine $\text{cov}(A_1'\hat{\beta}, A_2'\hat{\beta})$.

$$D = A'(X'X)^{-1}A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 7/3 & -1 \\ -1 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 35/6 \end{pmatrix} .$$

Hence, $d_{11} = 1/3$; $d_{12} = 2/3$; and $d_{22} = 35/6$. Now $P_1 - A_1'\hat{\beta} = 5 - 2 = 3$ and $P_2 - A_2'\hat{\beta} = 3 - 5/2 = 1/2$ so that

$$\frac{P_1 - A_1'\hat{\beta}}{P_2 - A_2'\hat{\beta}} = \frac{3}{1/2} = 6 .$$

Also

$$\frac{d_{12}}{d_{22}} = \frac{2/3}{35/6} = \frac{4}{35} .$$

Since $\frac{4}{35} \leq 6$, then

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} A_2 \left(\frac{P_2 - A_2'\hat{\beta}}{d_{22}} \right)$$

$$= \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} + \begin{pmatrix} 7/3 & -1 \\ -1 & 1/2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{1/2}{35/6} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{46}{35} \\ \\ \frac{13}{35} \end{pmatrix} \approx \begin{pmatrix} 1.314 \\ \\ .371 \end{pmatrix} .$$

Notice that

$$\tilde{\beta}_0 + 2\tilde{\beta}_1 = 2.056$$

and

$$2\tilde{\beta}_0 + \tilde{\beta}_1 = 3.000 ,$$

so that $\tilde{\beta}$ does satisfy the interval constraints.

Example 2

Assume a manufacturing department has 8 machines that can produce a certain item. At any given time the number of machines in operation can vary from 0 to 8. The number of defective items (per hour) were recorded when various numbers of machines were in operation and appear below. Assume it is known from past experience that when only one machine is in operation that the number of defective parts (per hour) never exceeds two.

Machines in Operation	0	1	2	3	4	5	6	7	8
Number of Defective Parts	0	2	5	7	10	12	13	15	17

For the model $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ $i = 0, 1, \ldots, 8$ where Y_i is the number of defective parts, X_i is the number of machines in operation, and the error terms are independent, we wish to find the least squares estimate for β that satisfies the preceding restriction. i.e. $0 \leq \beta_0 + \beta_1 \leq 2$.

For the given data,

$$(X'X)^{-1} = \frac{1}{540}\begin{pmatrix} 204 & -36 \\ -36 & 9 \end{pmatrix}; \quad \hat{\beta} = \begin{pmatrix} .467 \\ 2.133 \end{pmatrix};$$

and $A' = (1 \quad 1)$. Since $A'\hat{\beta} = 2.6 > 2$, ($\hat{\beta}$ violates the upper bound restriction) then from Table 3.1,

$$\tilde{\beta} = \hat{\beta} - \frac{(X'X)^{-1} A(A'\hat{\beta} - P_1)}{A'(X'X)^{-1}A}$$

$$= \begin{pmatrix} -.238 \\ 2.238 \end{pmatrix}.$$

CHAPTER IV. FIRST AND SECOND MOMENTS OF CONSTRAINED REGRESSION ESTIMATORS

Introduction

In the previous two chapters we derived the constrained regression estimators for β when β is restricted by linear inequalities or interval constraints. In this chapter we will now investigate the bias and mean squared error of these constrained estimators when A' is a $1 \times k$ matrix. We conclude this chapter with tables of these moments; the tables illustrate the fact that often the biased constrained estimator has a lower mean squared error than the unbiased unconstrained estimator.

Distribution of the Constrained Estimators

For the linear regression model $\underset{\sim}{Y} = \underset{\sim}{X\beta} + \epsilon$, where the error vector ϵ is multivariate normally distributed with mean vector 0 and variance-covariance matrix $\sigma_e^2 I$, then the unconstrained estimator for β , $\hat{\beta} = (X'X)^{-1}X'Y$, is also multivariate normally distributed with mean vector β and variance-covariance matrix $\sigma_e^2(X'X)^{-1}$.

If we introduce a linear inequality constraint on the β vector , i.e., $A'\beta \geq P$, then from Table 2.1, our estimator $\tilde{\beta}^{(1)}$ is given by:

$$\tilde{\beta}^{(1)} = \begin{cases} \hat{\beta} & \text{if} \quad A'\hat{\beta} \geq P \\ \hat{\beta} + \dfrac{(X'X)^{-1} A(P - A'\hat{\beta})}{A'(X'X)^{-1} A} & \text{if} \quad A'\hat{\beta} < P \ . \end{cases} \qquad (4.1)$$

Similarly if we impose a linear interval constraint on β, i.e., $P_0 \leq A'\beta \leq P_1$, then from Table 3.1, our estimator $\tilde{\beta}^{(2)}$ is given by:

$$\tilde{\beta}^{(2)} = \begin{cases} \hat{\beta} + \dfrac{(X'X)^{-1} A(P_0 - A'\hat{\beta})}{A'(X'X)^{-1} A} & \text{if} \quad A'\hat{\beta} < P_0 \\ \hat{\beta} & \text{if} \quad P_0 \leq A'\hat{\beta} \leq P_1 . \\ \hat{\beta} - \dfrac{(X'X)^{-1} A(A'\hat{\beta} - P_1)}{A'(X'X)^{-1} A} & \text{if} \quad A'\hat{\beta} > P_1 \end{cases} \tag{4.2}$$

Now the one-dimensional random variable $A'\hat{\beta}$ is normally distributed with mean $A'\beta$ and variance $\sigma_A^2 = \sigma_e^2 A'(X'X)^{-1}A$, so that the standardized variable $\dfrac{A'\hat{\beta} - A'\beta}{\sigma_A}$ is normally distributed with mean zero and variance one. From expressions (4.1) and (4.2), then

$$\frac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A} = \begin{cases} \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} & \text{if} \quad A'\hat{\beta} \geq P \\ \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} + \dfrac{A'(X'X)^{-1} A(P - A'\hat{\beta})}{\sigma_A A'(X'X)^{-1} A} & \text{if} \quad A'\hat{\beta} < P \end{cases}$$

$$= \begin{cases} \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} & \text{if} \quad \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} \geq \dfrac{P - A'\beta}{\sigma_A} \\ \dfrac{P - A'\beta}{\sigma_A} & \text{if} \quad \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} < \dfrac{P - A'\beta}{\sigma_A} \end{cases} \tag{4.3}$$

and

$$\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A} = \begin{cases} \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} + \dfrac{A'(X'X)^{-1}A(P_0 - A'\hat{\beta})}{\sigma_A A'(X'X)^{-1} A} & \text{if } A'\hat{\beta} < P_0 \\[2ex] \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} & \text{if } P_0 \leq A'\hat{\beta} \leq P_1 \\[2ex] \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} - \dfrac{A'(X'X)^{-1}A(A'\hat{\beta} - P_1)}{\sigma_A A'(X'X)^{-1} A} & \text{if } A'\hat{\beta} > P_1 \end{cases}$$

$$= \begin{cases} \dfrac{P_0 - A'\beta}{\sigma_A} & \text{if } \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} < \dfrac{P_0 - A'\beta}{\sigma_A} \\[2ex] \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} & \text{if } \dfrac{P_0 - A'\beta}{\sigma_A} \leq \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} \leq \dfrac{P_1 - A'\beta}{\sigma_A} \, . \\[2ex] \dfrac{P_1 - A'\hat{\beta}}{\sigma_A} & \text{if } \dfrac{A'\hat{\beta} - A'\beta}{\sigma_A} > \dfrac{P_1 - A'\hat{\beta}}{\sigma_A} \end{cases} \qquad (4.4)$$

The two estimators $\frac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}$ and $\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}$ have distributions that are partly continuous and partly discrete. The continuous part can be viewed as a truncated standard normal distribution and the probability associated with the discrete point or points can be evaluated in terms of areas under the tail or tails of the standard normal density function.

Moments of the Constrained Estimators

Define $\omega = \dfrac{A'\beta - P}{\sigma_A}$, $x = \dfrac{A'\beta - P_0}{\sigma_A}$, $y = \dfrac{P_1 - A'\beta}{\sigma_A}$,

$u = \left(\dfrac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2$, $F(k) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{k} e^{-\frac{1}{2}t^2} dt$ and $f(k) = \left(\dfrac{dF}{dt}\right)_{t=k} =$

$\dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}k^2}$.

First moments

$$E\left(\frac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}\right) = \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right) \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} \geq \frac{P - A'\beta}{\sigma_A}\right)$$

$$+ \left(\frac{P - A'\beta}{\sigma_A}\right) \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} < \frac{P - A'\beta}{\sigma_A}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{P-A'\beta}{\sigma_A}}^{\infty} \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right) e^{-\frac{1}{2}\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)$$

$$+ \left(\frac{P - A'\beta}{\sigma_A}\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{P-A'\beta}{\sigma_A}} e^{-\frac{1}{2}\left(\frac{A'\hat{\beta}-A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right) \tag{4.5}$$

Now the first function to be integrated in (4.5) is odd, so that the limits of integration can be changed to $\frac{A'\beta - P}{\sigma_A}$ and ∞ . Hence,

since

$$\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right) d \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right) = d\,(\tfrac{1}{2}u)$$

then

$$E\left(\frac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}\right) = \frac{1}{\sqrt{2\pi}} \int_{\omega^2}^{\infty} e^{-\frac{1}{2}u}\, d(\tfrac{1}{2}u) + -\omega F(-\omega)$$

$$= f(\omega) - \omega\, F(-\omega) \, . \qquad (4.6)$$

$$E\left(\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}\right) = \left(\frac{P_0 - A'\beta}{\sigma_A}\right) \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} < \frac{P_0 - A'\beta}{\sigma_A}\right)$$

$$+ \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right) \Pr\left(\frac{P_0 - A'\beta}{\sigma_A} \leq \frac{A'\hat{\beta} - A'\beta}{\sigma_A} \leq \frac{P_1 - A'\hat{\beta}}{\sigma_A}\right)$$

$$+ \left(\frac{P_1 - A'\hat{\beta}}{\sigma_A}\right) \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} > \frac{P_1 - A'\hat{\beta}}{\sigma_A}\right)$$

$$= \left(\frac{P_0 - A'\beta}{\sigma_A}\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{P_0 - A'\beta}{\sigma_A}} e^{-\frac{1}{2}\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)$$

$$+ \frac{1}{\sqrt{2\pi}} \int_{\frac{P_0 - A'\beta}{\sigma_A}}^{\frac{P_1 - A'\beta}{\sigma_A}} \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} \right) e^{-\frac{1}{2}\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} \right)$$

$$+ \left(\frac{P_1 - A'\hat{\beta}}{\sigma_A} \right) \frac{1}{\sqrt{2\pi}} \int_{\frac{P_1 - A'\hat{\beta}}{\sigma_A}}^{\infty} e^{-\frac{1}{2}\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} \right) . \tag{4.7}$$

Evaluating the middle integral of (4.7), we have

$$\frac{1}{\sqrt{2\pi}} \int_{(-x)^2}^{y^2} e^{-\frac{1}{2}u} \, d(\tfrac{1}{2}u) = \frac{1}{\sqrt{2\pi}} \int_{x^2}^{y^2} e^{-\frac{1}{2}u} \, d(\tfrac{1}{2}u)$$

$$= - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u} \Big|_{x^2}^{y^2}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} .$$

Therefore

$$E\left(\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A} \right) = - F(-x) + f(x) - f(y) + yF(-y) . \tag{4.8}$$

Second moments

$$E\left(\frac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}\right)^2 = \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2 \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} \geq \frac{P - A'\beta}{\sigma_A}\right)$$

$$+ \left(\frac{P - A'\beta}{\sigma_A}\right)^2 \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} < \frac{P - A'\beta}{\sigma_A}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{P-A'\beta}{\sigma_A}}^{\infty} \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2 e^{-\frac{1}{2}\left(\frac{A'\hat{\beta}-A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)$$

$$+ \left(\frac{P - A'\beta}{\sigma_A}\right)^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{P-A'\beta}{\sigma_A}} e^{-\frac{1}{2}\left(\frac{A'\hat{\beta}-A'\beta}{\sigma_A}\right)^2} d\left(\frac{A'\hat{\beta}-A'\beta}{\sigma_A}\right)$$

(4.9)

$$= \text{First integral of } (4.9) + \omega^2 F(-\omega) .$$

Now the first integral of (4.9) can be rewritten as

$$\frac{1}{\sqrt{2\pi}} \int_0^{\infty} t^2 e^{-\frac{1}{2}t^2} dt + \frac{1}{\sqrt{2\pi}} \int_0^{\omega} t^2 e^{-\frac{1}{2}t^2} dt ,$$

where

$$u = t^2 = \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2 . \qquad (4.10)$$

Evaluating the first integral of (4.10) , it follows that

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty t^2 e^{-\frac{1}{2}t^2} dt = \frac{1}{\sqrt{2\pi}} \int_0^\infty v^{\frac{1}{2}} e^{-v} dv ,$$

where

$$v = \tfrac{1}{2}\mu = \tfrac{1}{2}t^2 ,$$

which is equal to $\left(\frac{1}{\sqrt{\pi}}\right)\Gamma\left(\frac{3}{2}\right) = \frac{1}{\sqrt{\pi}}\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}$.

The final integral of (4.10) can be evaluated using tables of the incomplete gamma function or the normal distribution. In order to use the normal tables we outline the following development of Laning and Battin (1956, page 363). Let

$$H(z) = F(z) - F(0) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{1}{2}m^2} dm . \tag{4.11}$$

Substitute $m = t\sqrt{h}$ in expression (4.11) to obtain

$$\frac{1}{\sqrt{h}} H(z) = \frac{1}{\sqrt{2\pi}} \int_0^{z/\sqrt{h}} e^{-\frac{1}{2}t^2 h} dt . \tag{4.12}$$

Differentiating both sides of (4.12) with respect to h , we see that

$$\frac{H(z)}{h^{3/2}} = \frac{1}{\sqrt{2\pi}} \int_0^{z/\sqrt{h}} t^2 e^{-\frac{1}{2}t^2 h} dt + \frac{z}{h^{3/2}\sqrt{2\pi}} e^{-\frac{1}{2}z^2} . \tag{4.13}$$

After setting $h = 1$ and rearranging terms, equation (4.13) can be expressed as

$$\frac{1}{\sqrt{2\pi}} \int_0^z t^2 e^{-\frac{1}{2}t^2} dt = H(z) - \frac{z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} . \qquad (4.14)$$

Therefore, (4.10) is equal to $\frac{1}{2} + H(\omega) - \frac{\omega}{\sqrt{2\pi}} e^{-\frac{1}{2}\omega^2}$, where

$$H(\omega) = F(\omega) - F(0) = F(\omega) - \frac{1}{2} .$$

Consequently,

$$E\left(\frac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}\right)^2 = F(\omega) - \frac{\omega}{\sqrt{2\pi}} e^{-\frac{1}{2}\omega^2} + \omega^2 F(-\omega)$$

$$= F(\omega) + \omega^2 F(-\omega) - \omega f(\omega) . \qquad (4.15)$$

$$E\left(\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}\right)^2 = \left(\frac{P_0 - A'\beta}{\sigma_A}\right)^2 \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} < \frac{P_0 - A'\beta}{\sigma_A}\right)$$

$$+ \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2 \Pr\left(\frac{P_0 - A'\beta}{\sigma_A} \leq \frac{A'\hat{\beta} - A'\beta}{\sigma_A} \leq \frac{P_1 - A'\hat{\beta}}{\sigma_A}\right)$$

$$+ \left(\frac{P_1 - A'\hat{\beta}}{\sigma_A}\right)^2 \Pr\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A} > \frac{P_1 - A'\hat{\beta}}{\sigma_A}\right)$$

(4.16)

$$= x^2 F(-x) + \text{middle term of (4.16)} + y^2 F(-y) .$$

Evaluating the middle term of (4.16),

$$\left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2 \Pr\left(\frac{P_0 - A'\beta}{\sigma_A} \le \frac{A'\hat{\beta} - A'\beta}{\sigma_A} \le \frac{P_1 - A'\hat{\beta}}{\sigma_A}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-x}^{y} t^2 e^{-\frac{1}{2}t^2}\, dt\,, \quad \text{where} \quad t^2 = \left(\frac{A'\hat{\beta} - A'\beta}{\sigma_A}\right)^2,$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-x}^{0} t^2 e^{-\frac{1}{2}t^2}\, dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{y} t^2 e^{-\frac{1}{2}t^2}\, dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{x} t^2 e^{-\frac{1}{2}t^2}\, dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{y} t^2 e^{-\frac{1}{2}t^2}\, dt$$

$$= H(x) - xf(x) + H(y) - yf(y)$$

$$= F(x) - xf(x) + F(y) - yf(y) - 1\,.$$

Consequently ,

$$E\left(\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}\right)^2 = x^2F(-x) + F(x) - xf(x)$$

$$+ y^2F(-y) + F(y) - yf(y) - 1\,. \qquad (4.17)$$

Notice that as the constraints are relaxed the moment expressions given by (4.6), (4.8), (4.15) and (4.17) approach those of the unconstrained estimator. As $P \longrightarrow -\infty$, $P_0 \longrightarrow -\infty$ and $P_1 \longrightarrow +\infty$, then $\omega \longrightarrow \infty$, $x \longrightarrow \infty$ and $y \longrightarrow \infty$, so that (4.6) $\longrightarrow 0$, (4.8) $\longrightarrow 0$, (4.15) $\longrightarrow 1$ and (4.17) $\longrightarrow 1$.

Tables of Moments

In Table 4.1 we present the calculated moments of $\dfrac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}$ defined in (4.3), where $\sigma_A^2 = \sigma_e^2\, A'(X'X)^{-1} A$. Notice that for large values of $\omega = \dfrac{A'\beta - P_0}{\sigma_A}$, the bias is close to zero and the second moment is close to one. For smaller values of ω the effects of imposing the bound $A'\beta \geq P_0$ are more pronounced. However, the second moment is considerably less than that of the unconstrained estimator for many values of ω, including some negative values of ω. The minimum of the second moment occurs when $\omega = 0$, which corresponds to $A'\beta = P_0$. Consequently, the expected loss, as measured by the mean squared error, is reduced for a wide range of ω values once the bound is imposed. Only when ω is less than $-.75$, i.e., the bound is placed rather far on the wrong side of $A'\beta$, is the expected loss of the constrained estimator greater than that of the corresponding unconstrained estimator.

For the estimator constrained on both sides, $A'\tilde{\beta}^{(2)}$, the first and second moments of $\dfrac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}$ are tabled for various values of x and y in Tables 4.2 and 4.3. When the bounds are placed symmetrically with respect to $A'\beta$, the bias of $\dfrac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}$ is zero. As in the case of a single inequality constraint, for many values of x and y, the value of the mean squared error is substantially lower than that of the corresponding unconstrained estimator.

Table 4.1. Bias and mean squared error of $\dfrac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}$ when β is subject to the single constraint $A'\beta \geq P_0$

$\dfrac{A'\beta - P_0}{\sigma_A}$	$E\left(\dfrac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}\right)$	$E\left(\dfrac{A'\tilde{\beta}^{(1)} - A'\beta}{\sigma_A}\right)^2$
3.0000	0.0004	0.9975
2.7500	0.0009	0.9945
2.5000	0.0020	0.9888
2.2500	0.0042	0.9782
2.0000	0.0085	0.9603
1.7500	0.0162	0.9316
1.5000	0.0293	0.8892
1.2500	0.0506	0.8311
1.0000	0.0833	0.7580
0.7500	0.1312	0.6750
0.5000	0.1978	0.5926
0.2500	0.2863	0.5271
0.0000	0.3989	0.5000
-0.2500	0.5363	0.5354
-0.5000	0.6978	0.6574
-0.7500	0.8812	0.8875
-1.0000	1.0833	1.2420
-1.2500	1.3006	1.7314
-1.5000	1.5293	2.3608
-1.7500	1.7662	3.1309
-2.0000	2.0085	4.0397
-2.2500	2.2542	5.0843
-2.5000	2.5020	6.2612
-2.7500	2.7509	7.5680
-3.0000	3.0004	9.0025

Table 4.2. Bias of $\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}$ when β is subject to the single interval constraint $P_0 \leq A'\beta \leq P_1$

y \ x	3.0000	2.7500	2.5000	2.2500	2.0000	1.7500	1.5000	1.2500
3.0000	0.0	0.0005	0.0016	0.0039	0.0081	0.0158	0.0289	0.0502
2.7500	-0.0005	0.0	0.0011	0.0033	0.0076	0.0153	0.0284	0.0497
2.5000	-0.0016	-0.0011	0.0	0.0022	0.0065	0.0142	0.0273	0.0486
2.2500	-0.0039	-0.0033	-0.0022	0.0	0.0043	0.0119	0.0251	0.0464
2.0000	-0.0081	-0.0076	-0.0065	-0.0043	0.0	0.0077	0.0208	0.0421
1.7500	-0.0158	-0.0153	-0.0142	-0.0119	-0.0077	0.0	0.0131	0.0344
1.5000	-0.0289	-0.0284	-0.0273	-0.0251	-0.0208	-0.0131	0.0	0.0213
1.2500	-0.0502	-0.0497	-0.0486	-0.0464	-0.0421	-0.0344	-0.0213	0.0
1.0000	-0.0829	-0.0824	-0.0813	-0.0791	-0.0748	-0.0671	-0.0540	-0.0327
0.7500	-0.1308	-0.1303	-0.1292	-0.1269	-0.1227	-0.1150	-0.1019	-0.0806
0.5000	-0.1974	-0.1969	-0.1958	-0.1936	-0.1893	-0.1816	-0.1685	-0.1472
0.2500	-0.2860	-0.2854	-0.2843	-0.2821	-0.2779	-0.2702	-0.2570	-0.2358
0.0	-0.3986	-0.3484	-0.3980	-0.3969	-0.3947	-0.3905	-0.3828	-0.3696
-0.2500	-0.5360	-0.4858	-0.5354	-0.5343	-0.5321	-0.5279	-0.5202	-0.5070
-0.5000	-0.6974	-0.6472	-0.6969	-0.6958	-0.6936	-0.6893	-0.6816	-0.6685
-0.7500	-0.8808	-0.8306	-0.8803	-0.8792	-0.8769	-0.8727	-0.8650	-0.8519
-1.0000	-1.0829	-1.0327	-1.0824	-1.0813	-1.0791	-1.0748	-1.0671	-1.0540
-1.2500	-1.3002	-1.2500	-1.2997	-1.2986	-1.2964	-1.2921	-1.2844	-1.2713
-1.5000	-1.5289	-1.4787	-1.5284	-1.5273	-1.5251	-1.5208	-1.5131	-1.5000
-1.7500	-1.7658	-1.7156	-1.7653	-1.7642	-1.7619	-1.7577	-1.7500	-1.7369
-2.0000	-2.0081	-1.9579	-2.0076	-2.0065	-2.0043	-2.0000	-1.9923	-1.9792
-2.2500	-2.2539	-2.2036	-2.2533	-2.2522	-2.2500	-2.2457	-2.2381	-2.2249
-2.5000	-2.5016	-2.4514	-2.5011	-2.5000	-2.4978	-2.4935	-2.4858	-2.4727
-2.7500	-2.7505	-2.7003	-2.7500	-2.7489	-2.7467	-2.7424	-2.7347	-2.7216
-3.0000	-3.0000	-2.9498	-2.9995	-2.9984	-2.9961	-2.9919	-2.9842	-2.9711

$x = \frac{A'\beta - P_0}{\sigma_A}$ $y = \frac{P_1 - A'\beta}{\sigma_A}$

Table 4.2. (Continued)

y \ x	1.0000	0.7500	0.5000	0.2500	0.0	-0.2500	-0.5000	-0.7500	-1.0000
3.0000	0.0829	0.1308	0.1974	0.2860	0.3986	0.5360	0.6974	0.8808	1.0829
2.7500	0.0824	0.1303	0.1969	0.2854	0.3980	0.5354	0.6969	0.8803	1.0824
2.5000	0.0813	0.1292	0.1958	0.2843	0.3969	0.5343	0.6958	0.8792	1.0813
2.2500	0.0791	0.1269	0.1936	0.2821	0.3947	0.5321	0.6936	0.8769	1.0791
2.0000	0.0748	0.1227	0.1893	0.2779	0.3905	0.5279	0.6893	0.8727	1.0748
1.7500	0.0671	0.1150	0.1816	0.2702	0.3828	0.5202	0.6816	0.8650	1.0671
1.5000	0.0540	0.1019	0.1685	0.2570	0.3696	0.5070	0.6685	0.8519	1.0540
1.2500	0.0327	0.0806	0.1472	0.2358	0.3484	0.4858	0.6472	0.8306	1.0327
1.0000	0.0	0.0479	0.1145	0.2030	0.3156	0.4530	0.6145	0.7979	1.0000
0.7500	-0.0479	0.0	0.0666	0.1552	0.2678	0.4052	0.5666	0.7500	0.9521
0.5000	-0.1145	-0.0666	0.0	0.0885	0.2011	0.3385	0.5000	0.6834	0.8855
0.2500	-0.2030	-0.1552	-0.0885	0.0	0.1126	0.2500	0.4115	0.5948	0.7970
0.0	-0.3156	-0.2678	-0.2011	-0.1126	0.0	0.1374	0.2989	0.4822	0.6844
-0.2500	-0.4530	-0.4052	-0.3385	-0.2500	-0.1374	0.0	0.1615	0.3448	0.5470
-0.5000	-0.6145	-0.5666	-0.5000	-0.4115	-0.2989	-0.1615	0.0	0.1834	0.3855
-0.7500	-0.7979	-0.7500	-0.6834	-0.5948	-0.4822	-0.3448	-0.1834	0.0	0.2021
-1.0000	-1.0000	-0.9521	-0.8855	-0.7970	-0.6844	-0.5470	-0.3855	-0.2021	0.0
-1.2500	-1.2173	-1.1694	-1.1028	-1.0142	-0.9016	-0.7642	-0.6028	-0.4194	-0.2173
-1.5000	-1.4460	-1.3981	-1.3315	-1.2430	-1.1304	-0.9930	-0.8315	-0.6481	-0.4460
-1.7500	-1.6829	-1.6350	-1.5684	-1.4798	-1.3672	-1.2298	-1.0684	-0.8850	-0.6829
-2.0000	-1.9252	-1.8773	-1.8107	-1.7221	-1.6095	-1.4721	-1.3107	-1.1273	-0.9252
-2.2500	-2.1709	-2.1231	-2.0564	-1.9679	-1.8553	-1.7179	-1.5564	-1.3731	-1.1709
-2.5000	-2.4187	-2.3708	-2.3042	-2.2157	-2.1031	-1.9657	-1.8042	-1.6208	-1.4187
-2.7500	-2.6676	-2.6197	-2.5531	-2.4646	-2.3520	-2.2146	-2.0531	-1.8697	-1.6676
-3.0000	-2.9171	-2.8692	-2.8026	-2.7140	-2.6014	-2.4640	-2.3026	-2.1192	-1.9171

Table 4.2. (Continued)

y \ x	-1.2500	-1.5000	-1.7500	-2.0000	-2.2500	-2.5000	-2.7500	-3.0000
3.0000	1.3002	1.5289	1.7658	2.0081	2.2539	2.5016	2.7505	3.0000
2.7500	1.2997	1.5284	1.7653	2.0076	2.2533	2.5011	2.7500	2.9995
2.5000	1.2986	1.5273	1.7642	2.0065	2.2522	2.5000	2.7489	2.9984
2.2500	1.2964	1.5251	1.7619	2.0043	2.2500	2.4978	2.7467	2.9961
2.0000	1.2921	1.5208	1.7577	2.0000	2.2457	2.4935	2.7424	2.9919
1.7500	1.2844	1.5131	1.7500	1.9923	2.2381	2.4858	2.7347	2.9842
1.5000	1.2713	1.5000	1.7369	1.9792	2.2249	2.4727	2.7216	2.9711
1.2500	1.2500	1.4787	1.7156	1.9579	2.2036	2.4514	2.7003	2.9498
1.0000	1.2173	1.4460	1.6829	1.9252	2.1709	2.4187	2.6676	2.9171
0.7500	1.1694	1.3981	1.6350	1.8773	2.1231	2.3708	2.6197	2.8692
0.5000	1.1028	1.3315	1.5684	1.8107	2.0564	2.3042	2.5531	2.8026
0.2500	1.0142	1.2430	1.4798	1.7221	1.9679	2.2157	2.4646	2.7140
0.0	0.9016	1.1304	1.3672	1.6095	1.8553	2.1031	2.3520	2.6014
-0.2500	0.7642	0.9930	1.2298	1.4721	1.7179	1.9657	2.2146	2.4640
-0.5000	0.6028	0.8315	1.0684	1.3107	1.5564	1.8042	2.0531	2.3026
-0.7500	0.4194	0.6481	0.8850	1.1273	1.3731	1.6208	1.8697	2.1192
-1.0000	0.2173	0.4460	0.6829	0.9252	1.1709	1.4187	1.6676	1.9171
-1.2500	0.0	0.2287	0.4656	0.7079	0.9536	1.2014	1.4503	1.6998
-1.5000	-0.2287	0.0	0.2369	0.4792	0.7249	0.9727	1.2216	1.4711
-1.7500	-0.4656	-0.2369	0.0	0.2423	0.4881	0.7358	0.9847	1.2342
-2.0000	-0.7079	-0.4792	-0.2423	0.0	0.2457	0.4935	0.7424	0.9919
-2.2500	-0.9536	-0.7249	-0.4881	-0.2457	0.0	0.2478	0.4967	0.7461
-2.5000	-1.2014	-0.9727	-0.7358	-0.4935	-0.2478	0.0	0.2489	0.4984
-2.7500	-1.4503	-1.2216	-0.9847	-0.7424	-0.4967	-0.2489	0.0	0.2495
-3.0000	-1.6998	-1.4711	-1.2342	-0.9919	-0.7461	-0.4984	-0.2495	0.0

Table 4.3. Mean squared error of $\frac{A'\tilde{\beta}^{(2)} - A'\beta}{\sigma_A}$ when β is subject to the single interval constraint $P_0 \leq A'\beta \leq P_1$

y \ x	3.0000	2.7500	2.5000	2.2500	2.0000	1.7500	1.5000	1.2500
3.0000	0.9950	0.9921	0.9863	0.9758	0.9578	0.9291	0.8867	0.8286
2.7500	0.9921	0.9891	0.9833	0.9728	0.9548	0.9262	0.8838	0.8257
2.5000	0.9863	0.9833	0.9776	0.9670	0.9490	0.9204	0.8780	0.8199
2.2500	0.9758	0.9728	0.9670	0.9565	0.9385	0.9099	0.8675	0.8094
2.0000	0.9578	0.9548	0.9490	0.9385	0.9205	0.8919	0.8495	0.7914
1.7500	0.9291	0.9262	0.9204	0.9099	0.8919	0.8633	0.8209	0.7628
1.5000	0.8867	0.8838	0.8780	0.8675	0.8495	0.8209	0.7785	0.7203
1.2500	0.8286	0.8257	0.8199	0.8094	0.7914	0.7628	0.7203	0.6622
1.0000	0.7555	0.7526	0.7468	0.7363	0.7183	0.6897	0.6473	0.5891
0.7500	0.6725	0.6695	0.6638	0.6532	0.6353	0.6066	0.5642	0.5061
0.5000	0.5901	0.5871	0.5813	0.5708	0.5528	0.5242	0.4818	0.4237
0.2500	0.5246	0.5217	0.5159	0.5054	0.4874	0.4588	0.4164	0.3582
0.0	0.4975	0.4945	0.4888	0.4782	0.4603	0.4316	0.3892	0.3311
-0.2500	0.5329	0.5299	0.5242	0.5136	0.4956	0.4670	0.4246	0.3665
-0.5000	0.6549	0.6520	0.6462	0.6357	0.6177	0.5891	0.5467	0.4886
-0.7500	0.8850	0.8821	0.8763	0.8657	0.8478	0.8191	0.7767	0.7186
-1.0000	1.2395	1.2365	1.2307	1.2202	1.2022	1.1736	1.1312	1.0731
-1.2500	1.7289	1.7259	1.7202	1.7096	1.6916	1.6630	1.6206	1.5625
-1.5000	2.3583	2.3553	2.3495	2.3390	2.3210	2.2924	2.2500	2.1919
-1.7500	3.1284	3.1254	3.1196	3.1091	3.0911	3.0625	3.0201	2.9620
-2.0000	4.0372	4.0343	3.0285	4.0180	4.0000	3.9714	3.9290	3.8708
-2.2500	5.0818	5.0788	5.0730	5.0625	5.0445	5.0159	4.9735	4.9154
-2.5000	6.2587	6.2558	6.2500	6.2395	6.2215	6.1929	6.1504	6.0923
-2.7500	7.5655	7.5625	7.5567	7.5462	7.5282	7.4996	7.4572	7.3991
-3.0000	9.0000	8.9970	8.9913	8.9807	8.9628	8.9341	8.8917	8.8336

$$x = \frac{A'\beta - P_0}{\sigma_A} \qquad y = \frac{P_1 - A'\beta}{\sigma_A}$$

Table 4.3. (Continued)

y \ x	1.0000	0.7500	0.5000	0.2500	0.0	-0.2500	-0.5000	-0.7500	-1.0000
3.0000	0.7555	0.6725	0.5901	0.5246	0.4975	0.5329	0.6549	0.8850	1.2395
2.7500	0.7526	0.6695	0.5871	0.5217	0.4945	0.5299	0.6520	0.8821	1.2365
2.5000	0.7468	0.6638	0.5813	0.5159	0.4888	0.5242	0.6462	0.8763	1.2307
2.2500	0.7363	0.6532	0.5708	0.5054	0.4782	0.5136	0.6357	0.8657	1.2202
2.0000	0.7183	0.6353	0.5528	0.4874	0.4603	0.4956	0.6177	0.8478	1.2022
1.7500	0.6897	0.6066	0.5242	0.4588	0.4316	0.4670	0.5891	0.8191	1.1736
1.5000	0.6473	0.5642	0.4818	0.4164	0.3892	0.4246	0.5467	0.7767	1.1312
1.2500	0.5891	0.5061	0.4237	0.3582	0.3311	0.3665	0.4886	0.7186	1.0731
1.0000	0.5161	0.4330	0.3506	0.2851	0.2580	0.2934	0.4155	0.6455	1.0000
0.7500	0.4330	0.3500	0.2676	0.2021	0.1750	0.2104	0.3324	0.5625	0.9170
0.5000	0.3506	0.2676	0.1851	0.1197	0.0926	0.1279	0.2500	0.4801	0.8345
0.2500	0.2851	0.2021	0.1197	0.0542	0.0271	0.0625	0.1846	0.4146	0.7691
0.0	0.2580	0.1750	0.0926	0.0271	0.0000	0.0354	0.1574	0.3875	0.7420
-0.2500	0.2934	0.2104	0.1279	0.0625	0.0354	0.0708	0.1928	0.4229	0.7773
-0.5000	0.4155	0.3324	0.2500	0.1846	0.1574	0.1928	0.3149	0.5449	0.8994
-0.7500	0.6455	0.5625	0.4801	0.4146	0.3875	0.4229	0.5449	0.7750	1.1295
-1.0000	1.0000	0.9170	0.8345	0.7691	0.7420	0.7773	0.8994	1.1295	1.4839
-1.2500	1.4894	1.4064	1.3239	1.2585	1.2314	1.2668	1.3888	1.6189	1.9734
-1.5000	2.1188	2.0358	1.9533	1.8879	1.8608	1.8961	2.0182	2.2483	2.6027
-1.7500	2.8889	2.8059	2.7234	2.6580	2.6309	2.6662	2.7883	3.0184	3.3728
-2.0000	3.7978	3.7147	3.6323	3.5669	3.5397	3.5751	3.6972	3.9272	4.2817
-2.2500	4.8423	4.7592	4.6768	4.6114	4.5843	4.6196	4.7417	4.9718	5.3262
-2.5000	6.0192	5.9362	5.8538	5.7883	5.7612	5.7966	5.9187	6.1487	6.5032
-2.7500	7.3260	7.2429	7.1605	7.0951	7.0680	7.1033	7.2254	7.4555	7.8099
-3.0000	8.7605	8.6775	8.5951	8.5296	8.5025	8.5379	8.6599	8.8900	9.2445

Table 4.3. (Continued)

y \ x	-1.2500	-1.5000	-1.7500	-2.0000	-2.2500	-2.5000	-2.7500	-3.0000
3.0000	1.7289	2.3583	3.1284	4.0372	5.0818	6.2587	7.5655	9.0000
2.7500	1.7259	2.3553	3.1254	4.0343	5.0788	6.2558	7.5625	8.9970
2.5000	1.7202	2.3495	3.1196	4.0285	5.0730	6.2500	7.5567	8.9913
2.2500	1.7096	2.3390	3.1091	4.0180	5.0625	6.2395	7.5462	8.9807
2.0000	1.6916	2.3210	3.0911	4.0000	5.0445	6.2215	7.5282	8.9628
1.7500	1.6630	2.2924	3.0625	3.9714	5.0159	6.1929	7.4996	8.9341
1.5000	1.6206	2.2500	3.0201	3.9290	4.9735	6.1504	7.4572	8.8917
1.2500	1.5625	2.1919	2.9620	3.8708	4.9154	6.0923	7.3991	8.8336
1.0000	1.4894	2.1188	2.8889	3.7978	4.8423	6.0192	7.3260	8.7605
0.7500	1.4064	2.0358	2.8059	3.7147	4.7592	5.9362	7.2429	8.6775
0.5000	1.3239	1.9533	2.7234	3.6323	4.6768	5.8538	7.1605	8.5951
0.2500	1.2585	1.8879	2.6580	3.5669	4.6114	5.7883	7.0951	8.5296
0.0	1.2314	1.8608	2.6309	3.5397	4.5843	5.7612	7.0680	8.5025
-0.2500	1.2668	1.8961	2.6662	3.5751	4.6196	5.7966	7.1033	8.5379
-0.5000	1.3888	2.0182	2.7883	3.6972	4.7417	5.9187	7.2254	8.6599
-0.7500	1.6189	2.2483	3.0184	3.9272	4.9718	6.1487	7.4555	8.8900
-1.0000	1.9734	2.6027	3.3728	4.2817	5.3262	6.5032	7.8099	9.2445
-1.2500	2.4628	3.0921	3.8622	4.7711	5.8156	6.9926	8.2993	9.7339
-1.5000	3.0921	3.7215	4.4916	5.4005	6.4450	7.6220	8.9287	10.3633
-1.7500	3.8622	4.4916	5.2617	6.1706	7.2151	8.3921	9.6988	11.1334
-2.0000	4.7711	5.4005	6.1706	7.0795	8.1240	9.3009	10.6077	12.0422
-2.2500	5.8156	6.4450	7.2151	8.1240	9.1685	10.3455	11.6522	13.0867
-2.5000	6.9926	7.6220	8.3921	9.3009	10.3455	11.5224	12.8292	14.2637
-2.7500	8.2993	8.9287	9.6988	10.6077	11.6522	12.8292	14.1359	15.5704
-3.0000	9.7339	10.3633	11.1334	12.0422	13.0867	14.2637	15.5704	17.0050

CHAPTER V. GENERALIZED UPPER BOUNDING FOR QUADRATIC PROGRAMMING PROBLEMS

Introduction

Dantzig and Van Slyke (1967) developed a procedure to solve bounded linear programming problems with a specific set of size p orthogonal constraint equations. The general structure of such linear programming problems is as follows:

$$\text{minimize } x_0$$

subject to

$$m \text{ rows } \{ \overbrace{A^0 x_0 \cdots + A^{n_0} x_{n_0}}^{S_0} + \overbrace{\cdots\cdots\cdots\cdots\cdots}^{S_1} + \overbrace{A^{n_{p-1}+1} x_{n_{p-1}+1} \cdots + A^{n_p} x_{n_p}}^{S_p} = b \tag{5.1}$$

$$p \text{ rows } \left\{ \begin{array}{lll} x_{n_0+1} + \cdots x_{n_1} & & = 1 \\ & \ddots & \vdots \\ & x_{n_{p-1}+1} + \cdots x_{n_p} & = 1 \end{array} \right. \tag{5.2}$$

$$x_j \geq 0, \quad \text{all } j \;.$$

The columns A^j and the vector b have m components. Problems involving the delivery of goods from the supplier to the retailer, called transportation problems, are always of this form. Any linear programming problem with the last p equations orthogonal, and the right-hand side and all nonzero constraint coefficients positive, may be transformed into the above structure. To do so, we divide each equation by its right-hand side component and then scale the variables. Modifications for negative coefficients will be discussed in the section, "Adjustments to the Algorithm."

The procedure outlined here parallels Dantzig and Van Slyke's (1967) linear algorithm, in that it is basically a specialization of the simplex algorithm using a "working basis" of size $m \times m$ rather than of $(m+p) \times (m+p)$. When p is larger than m it is considerably cheaper to invert an $m \times m$ matrix than an $(m+p) \times (m+p)$ matrix. The following procedure requires much less core to store the information needed to carry out the calculations for each iteration than the traditional simplex algorithm.

In this chapter we allow the objective function to be quadratic, while retaining linearity in all constraint equations. Thus we make two kinds of modifications to Dantzig and Van Slyke's procedure: adjustments for a quadratic objective function and simplification for both linear and quadratic objective functions.

The algorithm will prove to be highly suitable for regression problems subject to side conditions.

The Structure of the Program

We first subdivide the columns into $p+1$ sets S_i, $i = 0, 1, \ldots, p$, such that each column in the set S_i, $i = 1, \ldots, p$, has a 1 as its

$(m+i)$th coefficient. The set S_0 consists of the columns with zeros as their $(m+1)$st through $(m+p)$th coefficients. Alternately we could think of dividing the variables up into the sets $S_0, \ldots, S_p$ as each variable x_j is associated with the column A^j. These $p+1$ sets are used to index the variables and columns.

Vectors with $m+p$ components are denoted by underlining A^j, i.e., $\underline{A}^j$. The corresponding vector containing only the first m components is not underlined. For example, the jth column of the system (5.1)-(5.2) is $\underline{A}^j$, and the first m components of the jth column is A^j.

For each set $S_1, \ldots, S_p$ we must choose a key column. The need for such key columns is a consequence of Theorem 1, to be presented in the section, "The Theory of the Algorithm." Since each of the last p rows sums to unity, the degrees of freedom for each row is one less than the number of variables in the row. Thus choosing one of the variables in the row, or alternately picking one column in each of the sets S_i, $i = 1, \ldots, p$, as "key" will collapse the subsystem (5.2) into a uniquely solvable system.

This key column can be selected in a variety of ways. Dantzig and Van Slyke (1967) select the first column in each set S_i, $i = 1, \ldots, p$, as the key, which allows for easier indexing. A rule for picking the optimal key columns is a possible tangent for further research. Until a better method is found, we suggest selecting from each set S_i, $i = 1, \ldots, p$, the column with the greatest cost coefficient. For a linear programming problem these columns certainly have the highest probability of entering the basis and are strong candidates for entering the basis of a quadratic program. If we select as key columns the vectors that will appear in the final basis, then we will eliminate iterations and calculations that would otherwise be made.

Notice that the set S_0 does not have a key column as it consists of the objective function, artificial and slack variables that must be added to the system (5.1)-(5.2), and any other variables not involved in the subsystem (5.2).

We will call the key columns $\underline{A}^{k_1}, \ldots, \underline{A}^{k_p}$, with $\underline{A}^{k_i}$ the key for set S_i .

The Theory of the Algorithm

The reduced system

Let us consider the following two theorems as given by Dantzig and Van Slyke (1967):

Theorem 1. Any feasible basis for (5.1)-(5.2) must include at least one column from each set S_i, $i = 0, 1, \ldots, p$.

Theorem 2. The number of sets containing two or more basic variables is at most $m-1$.

The importance of the key columns is a consequence of Theorem 1. Together the two theorems allow us to look at the structure of a feasible basis, for the program (5.1)-(5.2). Any feasible basis, $\underline{B}$, is of the form

$$\underline{B} = [\underline{A}^{j_1}, \underline{A}^{j_2}, \ldots, \underline{A}^{j_{m+p}}] \quad . \tag{5.3}$$

We assume that this feasible basis exists and is available. Finding an initial feasible basis is the goal of Phase 1 of the simplex method. Such a technique will be outlined in the section "Adjustments to the Algorithm."

We assume that this feasible basis $\underline{B} = [\underline{A}^{j_1}, \underline{A}^{j_2}, \ldots, \underline{A}^{j_{m+p}}]$ consists of all p key columns plus m additional columns. This is a reasonable assumption, since the keys will be rechosen whenever this situation fails to exist. The column $\underline{A}^0$ is always basic, and will occupy the $(m+p)$th column of every basis.

Since the key column of set S_i is denoted by $\underline{A}^{k_i}$, then the basis in (5.3) can be partitioned as shown in Figure 5.1.

I_{pxp} is an identity matrix of dimension p, and the m vectors in the submatrix C_{pxm} are either zero vectors (columns in set S_0) or unit vectors. The $(p+t)$th column of $\underline{B}$ is from the set S_{r_t}, $r_t = 0,1,\ldots,p$.

Due to the structure of the program, all columns in the same set have identical last p components. Furthermore, due to the specific organization of (5.2), each column in set S_i consists of a unit vector in the last p components, with the 1 appearing in the $(m+i)$th position. Thus, if in Figure 5.1, each of the keys $\underline{A}^{k_i}$ is subtracted from each of the other basic columns in its set, C_{pxm} is reduced to a matrix of zeros. This new matrix $\underline{BT}$ is as shown in (5.4).

$$\begin{array}{cc} \text{Key columns} & \text{Nonkey columns} \\ \underline{A}^{k_1}, \underline{A}^{k_2}, \ldots, \underline{A}^{k_p} & \text{tth column from Set } S_{r_t} \end{array}$$

$$\underline{B} = \left[\begin{array}{c|c} A_{mxp} & N_{mxm} \\ \hline I_{pxp} & C_{pxm} \end{array}\right] \begin{array}{l} \left.\vphantom{A_{mxp}}\right\} m \text{ rows} \\ \left.\vphantom{I_{pxp}}\right\} p \text{ rows} \end{array}$$

Figure 5.1. A partitioned feasible basis

$$\underline{BT} = \left[\begin{array}{c|c} A_{mxp} & B_{mxm} \\ \hline I_{pxp} & 0_{pxm} \end{array}\right] . \tag{5.4}$$

This matrix consists of two submatrices from $\underline{B}$: A_{mxp} and I_{pxp}. The 0_{pxm} submatrix classifies $\underline{BT}$ as an upper block triangular matrix, and B_{mxm} is the resulting matrix due to the changes to N_{mxm}.

The submatrix B_{mxm} is extremely important. Recall that $\underline{BT}$ is merely the result of subtracting from each nonkey basic column its corresponding key. Therefore, B_{mxm} is as follows:

$$B_{mxm} = [N_{r_1} - A^{k_{r_1}}, N_{r_2} - A^{k_{r_2}}, \ldots, N_{r_m} - A^{k_{r_m}}] .$$

Note that if $N_{r_i} \varepsilon S_0$, then $N_{r_i} - A^{k_{r_i}} = N_{r_i}$ since there is no key for S_0 .

Once again due to the specific structure of (5.2) and the ordering of the keys $A^{k_1} \ldots A^{k_p}$ in the submatrix A_{mxp}, the resulting matrix B_{mxm} can be formed through a matrix multiplication.

$A_{mxp}C_{pxm} = Q_{mxm}$, where Q_{mxm} is of the form:

$$Q_{mxm} = [A^{k_{r_1}}, A^{k_{r_2}}, A^{k_{r_3}}, \ldots, A^{k_{r_m}}] ,$$

with $A^{k_{r_i}}$ a column of zeros if $N_{r_i} \varepsilon S_0$.

Putting the above operations together, we have

$$B_{mxm} = N_{mxm} - Q_{mxm} = N_{mxm} - A_{mxp}C_{pxm} .$$

Expanding the matrix multiplication to all of $\underline{B}$, we find that the following matrix $\underline{T}$ is precisely the one needed:

$$\underline{T} = \left[\begin{array}{c|c} I_{pxp} & -C_{pxm} \\ \hline 0_{mxp} & I_{mxm} \end{array}\right] .$$

The det (T) = 1, so T is nonsingular.

Checking, we realize that, in fact,

$$\underline{BT} = \left[\begin{array}{c|c} A_{mxp} & N_{mxm} \\ \hline I_{pxp} & C_{pxm} \end{array}\right] \left[\begin{array}{c|c} I_{pxp} & -C_{pxm} \\ \hline 0_{mxp} & I_{mxm} \end{array}\right]$$

$$= \left[\begin{array}{c|c} A_{mxp} & N_{mxm} - A_{mxp}C_{pxn} \\ \hline I_{pxp} & 0_{pxm} \end{array}\right]$$

as defined by expression (5.4).

This has, in effect, collapsed the system (5.1)-(5.2) into one of size mxm by reducing the number of "working variables" from m+p to m. Since

$$\underline{B}\, x_{\underline{B}} = \underline{b}\ , \quad \text{where } x_B \text{ are the basic variables,}$$

then

$$(\underline{BT})\ (\underline{T}^{-1}\underline{x_B}) \ = \ \underline{b}\ .$$

Letting $\underline{T}^{-1}\underline{x_B} = \underline{y_B}$

then $\underline{BT}\ \underline{y_B} = \underline{b}\ .$

Expanding this equation,

$$\left[\begin{array}{c|c} A_{mxp} & N_{mxm} - A_{mxp}C_{pxm} \\ \hline I_{pxp} & 0_{pxm} \end{array}\right] \left[\begin{array}{c} y_1 \\ \vdots \\ y_p \\ \hline y_{p+1} \\ \vdots \\ y_{p+m} \end{array}\right] = \left[\begin{array}{c} b_1 \\ \vdots \\ b_m \\ \hline b_{m+1} \\ \vdots \\ b_{p+m} \end{array}\right] = \left[\begin{array}{c} b_1 \\ \vdots \\ b_m \\ \hline 1 \\ \vdots \\ 1 \end{array}\right] .$$

Multiplying,

$$\left[A_{mxp}\right]\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} + \left[N_{mxm}\right]\begin{bmatrix} y_{p+1} \\ \vdots \\ y_{p+m} \end{bmatrix} - \left[A_{mxp}C_{pxm}\right]\begin{bmatrix} y_{p+1} \\ \vdots \\ y_{p\ m} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix} \qquad (5.5)$$

and

$$\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} b_{m+1} \\ \vdots \\ b_{m+p} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} . \qquad (5.6)$$

If in (5.5), $\left[A_{mxp}\right]\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}$ is subtracted from both sides of the equation, then

$$\left[N_{mxm}\right]\begin{bmatrix} y_{p+1} \\ \vdots \\ y_{p+m} \end{bmatrix} - \left[A_{mxp}C_{pxm}\right]\begin{bmatrix} y_{p+1} \\ \vdots \\ y_{p+m} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix} - \left[A_{mxp}\right]\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} . \tag{5.7}$$

Substituting (5.6) into (5.7) and putting together the left-hand side of (5.7), then

$$\left[B_{mxm}\right]\begin{bmatrix} y_{p+1} \\ \vdots \\ y_{p+m} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix} - \left[A_{mxp}\right]\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{5.8}$$

$$= b - \sum_{i=1}^{p} A^{k_i} \equiv d \ .$$

Define $D^j = A^j - A^{k_i}$ if $A^j \in S_i$.

Then by the structure of B_{mxm} ,

$$B_{mxm} \equiv B = \{D^j \mid A^j \text{ is basic and not key}\} \ .$$

Using this definition for B in (5.8) results in the "reduced system" of size mxm:

$$\sum_{j=1}^{m} D^j y_{p+j} = d .$$

B is called the "working basis" and is clearly made up of columns from the reduced system.

The preceding presentation gives us the following result:

Theorem 3. The working basis, B, is a basis for the reduced system.

One complete iteration

Recall that $\underline{B}$ is a feasible basis for the program (5.1)-(5.2) and B is a basis for the corresponding reduced system.

Computing the simplex multipliers

Definition 5.1. The simplex multipliers $(\pi:\pi^*)$ where π is associated with the first m equations and π^* the last p , form a vector satisfying the following equation:

$$(\pi:\pi^*) = (\pi_1,\ldots,\pi_m:\pi^*_{m+1},\ldots,\pi^*_{m+p}) = \underline{c}_B\underline{B}^{-1} , \tag{5.9}$$

where $\underline{c}_B$ is the vector of cost coefficients for the basic variables.

For our program, all the ordinary variables that are defined in the quadratic system (5.1)-(5.2) have a cost coefficient of zero, except for x_0 which has a cost coefficient of (positive) one. Therefore,

$$\underline{c}_B = (0, 0, 0, \ldots, 0, 1) .$$

The cost coefficients for added artificial variables will be discussed in the section "Adjustments to the Algorithm." Multiplying (5.9) on the right by $\underline{B}$,

$$(\pi:\pi^*)\ \underline{B} = \underline{c}_B = (0, 0, 0, \ldots, 0, 1) \quad . \tag{5.10}$$

Multiplying (5.10) by $\underline{T}$,

$$(\pi:\pi^*)\ (\underline{BT}) = (0, 0, 0, \ldots, 0, 1)\underline{T} \tag{5.11}$$

$$= (0, 0, 0, \ldots, 0, 1) \left[\begin{array}{c|c} I_{pxp} & -C_{pxm} \\ \hline 0_{mxp} & I_{mxm} \end{array}\right]$$

$$= (0, 0, 0, \ldots, 0, 1) \ .$$

Using the partitioned form of $\underline{BT}$, (5.11) then becomes

$$(\pi:\pi^*) \left[\begin{array}{c|c} A_{mxp} & B \\ \hline I_{pxp} & 0_{pxm} \end{array}\right] = (0, 0, 0, \ldots, 0, 1) \ .$$

Multiplying,

$$\pi A^{k_i} + \pi_i^* = 0 \qquad i = 1, \ldots, p \tag{5.12}$$

and

$$\pi B = (\overbrace{0, 0, 0, \ldots, 0, 1}^{m \text{ components}}) . \tag{5.13}$$

Since B^{-1} exists, multiplying (5.13) by B^{-1},

$$\pi = (0, 0, 0, \ldots, 0, 1)B^{-1} ,$$

which implies that π is the mth row of B^{-1}. Subtracting πA^{k_i} from both sides of (5.12),

$$\pi_i^* = -\pi A^{k_i} , \qquad i = 1, \ldots, p . \tag{5.14}$$

Therefore if B^{-1} is known, then a full set of prices for the original system can be calculated.

<u>Determining the column to enter the basis</u> This is done in the usual simplex manner, by computing

$$c_j = (\pi : \pi^*)\underline{A}^j = \pi A^j + \pi_i^* , \qquad \text{if} \quad A^j \in S_i .$$

Substituting (5.14) into the above equation,

$$c_j = \pi A^j - \pi A^{k_i} ,$$

$$= \pi(A^j - A^{k_i}) , \qquad \text{if} \qquad A^j \in S_i , \tag{5.15}$$

for nonbasic columns A^j .

Keeping in mind the cross-product equations (i.e. $x_i\lambda_i = 0$ and $y_i s_i = 0$), we select as the entering variable the first (from left-to-right) possible variable whose corresponding c_j is negative. If no such variable exists, the current basic solution is either optimal or infeasible. Let $\underline{A}^s \in S_\sigma$ enter the basis.

<u>Finding the representation of the entering column in terms of the current basis</u> When $\underline{A}^s \in S_\sigma$ enters the basis of the entire system (5.1)-(5.2), then correspondingly $D^s = A^s - A^{k_\sigma}$ enters the basis of the reduced system. Following the revised simplex algorithm, A^s must be updated via B^{-1} before it enters the basis, i.e.,

$$\underline{A}^{s*} = \underline{B}^{-1}\underline{A}^s . \tag{5.16}$$

Likewise

$$D^{s*} = B^{-1}(A^s - A^{k_\sigma}) \tag{5.17}$$

is the current representation of $(A^s - A^{k_\sigma})$ that actually enters the basis of the reduced system. Let us investigate the relationship between $\underline{A}^{s*}$ and D^{s*}. We will again make use of the matrix $\underline{T}$, since B is the upper-right sub-matrix of $\underline{BT}$. Multiplying both sides of (5.16) by $\underline{B}$,

$$\underline{B}\,\underline{A}^{s*} = \underline{A}^s . \tag{5.18}$$

We make a transformation using the matrix $\underline{T}$:

$$\underline{A}^{s*} = \underline{TZ}^{s} , \tag{5.19}$$

so that (5.18) becomes

$$(\underline{BT})\underline{Z}^{s} = \underline{A}^{s} . \tag{5.20}$$

We will solve the above equation for Z^s .

In partitioned form, (5.20) is

$$\left[\begin{array}{c|c} A_{mxp} & B \\ \hline I_{pxp} & 0_{pxm} \end{array}\right] \begin{bmatrix} Z_1^s \\ \vdots \\ Z_p^s \\ \vdots \\ Z_{p+1}^s \\ \vdots \\ Z_{p+m}^s \end{bmatrix} = \left[\begin{array}{c} A_s \\ \hline 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{array}\right] \begin{array}{c} \\ \\ \\ \\ \leftarrow \text{position } \sigma . \\ \\ \\ \end{array}$$

The bottom p equations of (5.20) yield

$$Z_i^s = 0 , \qquad 1 \leq i \leq p \quad i \neq \sigma$$

and

$$Z_\sigma^s = 1 .$$

Using the above values in the top m equations of (5.20),

$$A^{k_\sigma} + B \begin{bmatrix} Z_{p+1}^s \\ \vdots \\ Z_{p+m}^s \end{bmatrix} = A^s . \tag{5.21}$$

Subtracting A^{k_σ} from both sides of (5.21),

$$B \begin{bmatrix} Z^s_{p+1} \\ \vdots \\ Z^s_{p+m} \end{bmatrix} = A^s - A^{k_\sigma} .$$

Multiplying both sides by B^{-1} ,

$$\begin{bmatrix} Z^s_{p+1} \\ \vdots \\ Z^s_{p+m} \end{bmatrix} = B^{-1}(A^s - A^{k_\sigma}) . \tag{5.22}$$

From (5.17) we know that $B^{-1}(A^s - A^{k_\sigma}) = D^{s*}$, so that the vector

$$\begin{bmatrix} Z^s_1 \\ \vdots \\ Z^s_p \\ Z^s_{p+1} \\ \vdots \\ Z^s_{p+m} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ \hline D^{s*}_1 \\ \vdots \\ D^{s*}_m \end{bmatrix} \begin{matrix} \\ \\ \longleftarrow \text{position } \sigma \\ \\ \\ \\ \\ \\ \end{matrix} .$$

In view of (5.19) ,

$$\underline{A}^{s*} = \begin{bmatrix} A_1^{s*} \\ \vdots \\ A_p^{s*} \\ \hline A_{p+1}^{s*} \\ \vdots \\ A_{p+m}^{s*} \end{bmatrix} = \left[\begin{array}{c|c} I_{pxp} & -C_{pxm} \\ \hline 0_{mxp} & I_{mxm} \end{array}\right] \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ \hline D_1^{s*} \\ \vdots \\ D_m^{s*} \end{bmatrix} \begin{matrix} \\ \\ \longleftarrow \text{position } \sigma \\ \\ \\ \\ \\ \\ \end{matrix} . \tag{5.23}$$

Therefore, A_i^{s*} is the scalar product of the ith row of $\underline{T}$ with the vector $\underline{Z}^s$. Recall that the tth column of $-C_{pxm}$ has a -1 in position r_t, (where column $p+t$ of $\underline{B}$ is a member of group S_{r_t}) and zeros everywhere else. Therefore, whenever $r_t = i$, the product of the ith row of $-C_{pxm}$ with D^{s*} will be nonzero.

Special attention should be drawn to the $(m+p)$th component of $\underline{A}^{s*}$. Recall that the mth row of B^{-1} consists of the vector π. Since $\underline{A}^s$ is entering the basis, then

$$c_s = \pi(A^s - A^{k_\sigma}) = B^{-1}_{mth\ row}(A^s - A^{k_\sigma}) < 0 .$$

This is precisely the value of $\underline{A}^{s*}_{m+p}$, because

$$\underline{A}^{s*}_{m+p} = D_m^{s*} = (B^{-1})_m (A^s - A^{k_\sigma}) .$$

<u>Choosing the column to leave the basis</u> This is done in the usual simplex way by computing

$$\min_{A_i^{s*} > 0} \left\{ \frac{b_i}{A_i^{s*}} \right\} = \frac{b_r}{A_r^{s*}} = \theta \ , \tag{5.24}$$

where b_i is the current value for the ith basic variable. If all $A_i^{s*} \leq 0$, the program (5.1)-(5.2) has an unbounded solution. Otherwise, column r of $\underline{B}$, $\underline{A}^{j_r}$, leaves the basis. Let $\underline{A}^{j_r} \varepsilon S_\rho$.

<u>Updating the values of the basic variables</u> This is also done using the standard simplex formulae:

$$(b_i^*)_{new} = (b_i^*)_{old} - \theta A_i^{s*} \ , \quad i = 1,\ldots,m+p \quad i \neq r$$

and (5.25)

$$(b_r^*)_{new} = \theta \ .$$

<u>Updating the inverse of the working basis</u> Recall that the column entering the basis is $\underline{A}^s \varepsilon S_\sigma$ and the column leaving is $\underline{A}^{j_r} \varepsilon S_\rho$. There are two different cases we must consider.

Case 1: $\underline{A}^{j_r}$ is not a key column.

$\underline{A}^{j_r}$ is one of the last m columns of $\underline{B}$, say column $p+i_2$. When $\underline{A}^{j_r}$ is replaced in $\underline{B}$ by $\underline{A}^s$, then the corresponding change in B is that $A^{j_r} - A^{k_\rho}$ is replaced by $A^s - A^{k_\sigma}$. Thus to update B^{-1}, adjoin to it the column

$$D^{s*} = B^{-1}(A^s - A^{k_\sigma}) .$$

We then perform a pivot operation, so that the adjoined column becomes a unit vector with i_2 the unit component. Naturally, the vector π changes as B^{-1} does.

Case 2: $\underline{A}^{j_r}$ is a key column.

Subcase 2a: $\underline{A}^{j_r}$ and $\underline{A}^s$ are from different sets. Since $\rho \neq \sigma$, then we are replacing the key column of S_ρ with a column from S_σ. Therefore we must choose a new key column for S_ρ, because at least one column from S_ρ must be basic at all times. In view of Theorem 1, $\underline{B}_j$ must contain a nonkey column of S_ρ in addition to $\underline{A}^{j_r}$, or else $\underline{A}^{j_r}$ could not leave when $\underline{A}^s$ enters. As the new key for S_ρ we select any nonkey column of S_ρ that is presently a member of $\underline{B}$, and call it $\underline{A}^k$, where $k = p + i_2$. It will become key by interchanging it with $\underline{A}^{j_r}$ (which is also $\underline{A}^{k_\rho}$). When this change is made in $\underline{B}$, the corresponding change in B is as follows:

$$A^{j_i} - A^{k_\rho} \text{ is replaced with } A^{j_i} - A^k \text{ when } A_{j_i} \in S_\rho \quad j_i \neq k$$

and

$$A^k - A^{k_\rho} \text{ is replaced with } A^{k_\rho} - A^k .$$

These replacements can all be made by multiplying $A^k - A^{k_\rho}$ by -1 and then adding the results to the original columns of B, where $A^{j_i} - A^{k_\rho}$ with $A_{j_i} \in S_\rho$. In matrix form, B is replaced with BT_1, where

$$T_1 = \begin{bmatrix} 1 & & & & & & & & \\ \vdots & \ddots & & & & & & & \\ \vdots & & 1 & & & & & & \\ 0 & \cdots & -1 & \cdots 0 \cdots & -1 & \cdots & 0 & \cdots -1 \cdots & 0 \\ \vdots & & & & & & 1 & & \vdots \\ \vdots & & & & & & & \ddots & \vdots \\ \vdots & & & & & & & 1 & \vdots \\ & & & & & & & & \ddots \\ & & & & & & & & 1 \end{bmatrix} \longleftarrow \text{row } i_2 \; . \qquad (5.26)$$

The -1's occur in the columns corresponding to $\underline{A}^{j_i} \in S_\rho$. Recall that the ones in the ρth row of C also appear in these positions, so that row i_2 of T_1 is the negative of row ρ of C .

B^{-1} is thus replaced by $T_1^{-1}B^{-1}$. It is easy to see that $T_1 \equiv T_1^{-1}$, because if we used the same transformation T_1 twice, we would obtain the original matrix B . Therefore B^{-1} is replaced with T_1B^{-1} . Once we make this change, the problem is the same as that in Case 1.

Subcase 2b: $\rho = \sigma$. This time the key column of S_ρ is being replaced by another column from the same set. If there are other nonkey columns from S_ρ in $\underline{B}$, we interchange one of these with $\underline{A}^{j_r}$ as in subcase 2a. We then perform the operations of Case 1. However, if no nonkey column of S_ρ is a member of $\underline{B}$, then we replace $\underline{A}^{j_r}$ with $\underline{A}^s$ in $\underline{B}$ and it becomes the new key of S_ρ . Since the last m columns of $\underline{B}$ contain no column from S_ρ , then we do not change B^{-1} .

While performing subcase 2b, we are correcting for having not chosen the "best" key for S_ρ . This is a simple change to perform, but if A^s had been chosen as key for S_ρ, the iteration would have been avoided.

Algorithm flow chart

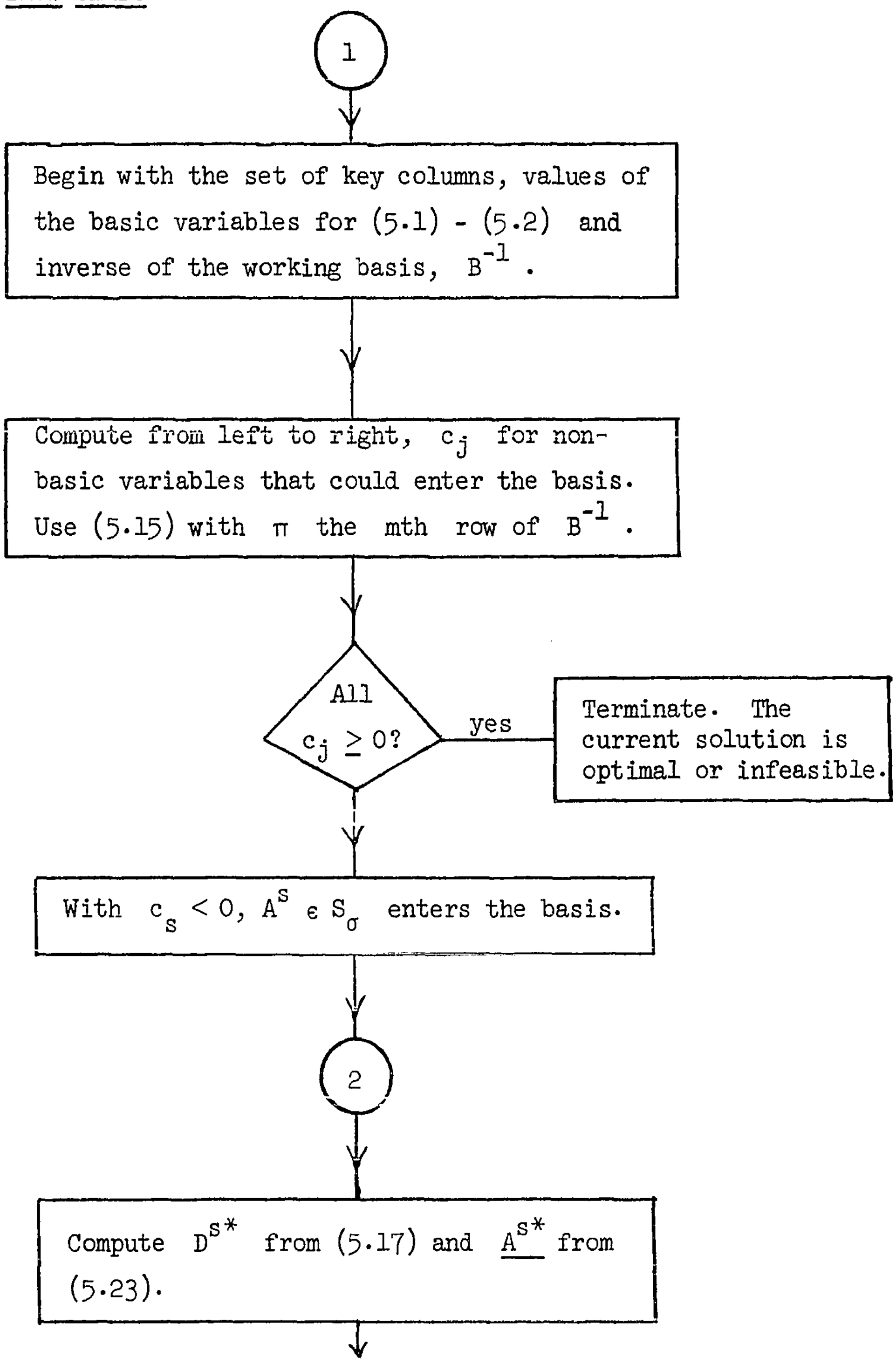

Figure 5.2. Flow chart for the generalized upper bounding algorithms.

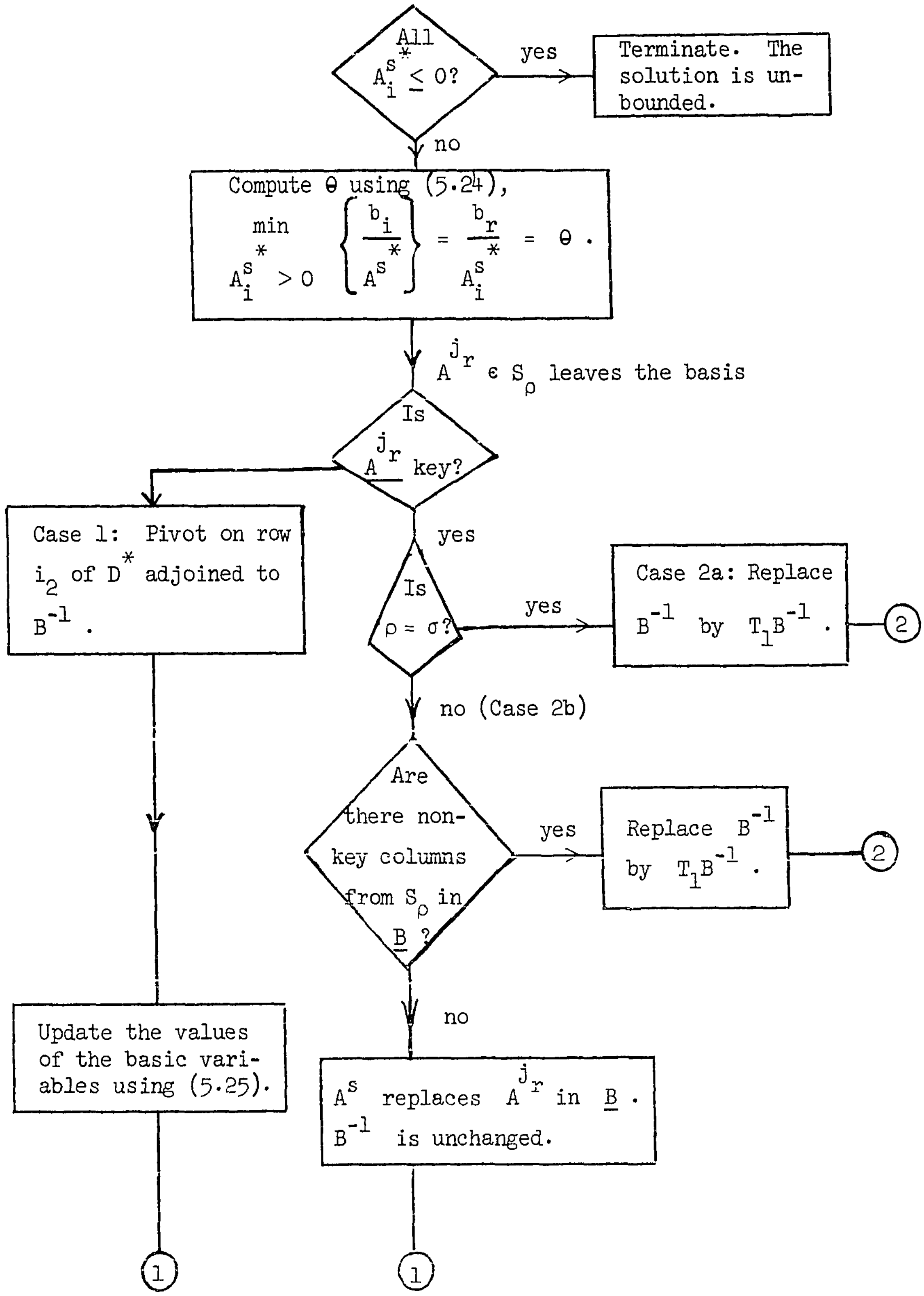

Figure 5.2. (Continued)

Adjustments to the Algorithm

Phase I Artificial variables

In order to arrive at an initial feasible solution, we can add artificial variables to the system (5.1)-(5.2) as in Phase I of the simplex algorithm. We assign cost coefficients of minus one (or any value less than zero), to insure that they will not remain in the basis. These artificials are added for convenience to begin the algorithm. We generally adjoin $m-1$ of them to the initial tableau, thus forming the columns of an $(m-1) \times (m-1)$ identity matrix.

Because of the cost coefficient values of -1, the objective function is temporarily changed from $x_0 = f(x)$ to $f(x) - \sum_{i=1}^{m-1} w_i$, where w_i, $i = 1, \ldots, m-1$ are the added artificial variables. Therefore, instead of maximizing x_0, we maximize $x_0 + \sum_{i=1}^{m-1} w_i$.

All of these w_i are members of S_0, so that the beginning basis is

$$[\underline{A}^{j_1}\ \underline{A}^{j_2} \ldots \underline{A}^{j_p} | w_1\ w_2 \ldots w_{m-1}\ \underline{A}^0] .$$

Once we delete all the w_i from the basis, we have an initial feasible basis. Due to the -1 cost coefficients these artificials will never re-enter the basis once they have left the basis.

<u>Modifications for negative coefficients in the last p equations and statistical applications</u>

After dividing each of the last p equations by its right-hand side and scaling the variables, we have right-hand side components with values of positive one, and the coefficients of the variables are either positive one or negative one. Since each set S_i must have at least one variable with a positive one coefficient, then we could assign the key columns so that all of them have plus one coefficients. When forming a "working basis," if a nonkey basic column has a -1 coefficient, we then add the key column to the basic matrix rather than subtract this column. Also equation (5.15) changes to $c_j = \pi(A^j + A^{k_i})$ if $A^j \varepsilon S_i$ and A^j has a -1 coefficient in the last p equations. Similar modifications must be made to the formulae necessary to update B^{-1} and compute A_i^{s*}.

The above situation occurs in the following regression problems with side conditions:

$$1)\quad \min(y - X\beta)'(y - X\beta) = y'y + \min[-2y'X\beta + \beta'X'X\beta]$$

$$\text{subject to } \sum_{i=1}^{n} \beta_i = 0 \, .$$

Then if we let $\beta_1^* = \beta_1 + 1$,

$$\sum_{i=1}^{n} \beta_i = 0 \text{ can be equivalently expressed as}$$

$$\beta_1^* + \beta_2 + \cdots + \beta_n = 1 \, .$$

If we then double the number of β variables so that

$$\beta_i = \beta_i^{(1)} - \beta_i^{(2)}; \quad \beta_i^{(1)} \geq 0; \quad \beta_i^{(2)} \geq 0 \quad \text{for all } i$$

then the entire problem can be rewritten as:

$$y'y + \min[-2y'X\beta + \beta'X'X\beta]$$

subject to

$$\beta_1^{*(1)} + \beta_2^{(1)} + \dots + \beta_n^{(1)} - (\beta_1^{*(2)} + \dots + \beta_n^{(2)}) = 1$$

$$\beta_i^{(j)} \geq 0 \qquad j = 1, 2 \text{ and } i = 1, 2, \dots, n .$$

2) $y'y + \min[-2y'X\beta + \beta'X'X\beta]$

subject to $\beta_1 \leq 1$, i.e., $\beta_1 \in (-\infty, 1]$.

Then if we again let $\beta_1 = \beta_1^{(1)} - \beta_1^{(2)}$; $\beta_1^{(1)}, \beta_1^{(2)} \geq 0$ and incorporate β_2^s as a slack variable, the problem becomes

$$y'y + \min[-2y'X\beta + \beta'X'X\beta]$$

subject to

$$\beta_1^{(1)} - \beta_1^{(2)} + \beta_2^s = 1$$

$$\beta_1^{(1)}, \beta_1^{(2)}, \beta_2^s \geq 0 .$$

Both of the above problems are quadratic programming problems with $y'y$ a constant, $-2y'X\beta$ a linear function in the β's , and $\beta'X'X\beta$ a quadratic form in the β's .

Example

Consider the problem:

$$\max \quad x_1 + x_2 - \tfrac{1}{2}x_1^2 + x_1x_2 - x_2^2$$

$$\text{subject to} \quad x_1 + x_2 \leq 1$$

$$2x_1 + 3x_2 \geq 2$$

$$x_1, x_2 \geq 0 \ .$$

<u>The original tableau</u>

S_0	S_1	S_1	S_1	S_0	S_0	S_0	S_0	S_0	S_0	S_0	S_0	<u>b</u>
x_0	x_1	x_2	s_2	s_1	y_1	y_2	λ_1	λ_2	w_1	w_2	w_3	
A^0	A^1	A^2	A^3	A^4	A^5	A^6	A^7	A^8	A^9	A^{10}	A^{11}	
1	0	0	0	0	0	0	0	0	1	1	1	-4
0	1	-1	0	0	-2	1	-1	0	1	0	0	1
0	-1	2	0	0	-3	1	0	-1	0	1	0	1
0	2	3	0	-1	0	0	0	0	0	0	1	2
0	1	1	1	0	0	0	0	0	0	0	0	1

$$\underline{B} = (A^3 | A^9 A^{10} A^{11} A^0) ; \qquad \underline{b} = (1,1,1,2,-4) .$$

$$\underline{B} = \left.\begin{bmatrix} 1 & | & 1 & 1 & 1 & 1 \\ 0 & | & 1 & 0 & 0 & 0 \\ 0 & | & 0 & 1 & 0 & 0 \\ 0 & | & 0 & 0 & 1 & 0 \\ \hline 1 & | & 0 & 0 & 0 & 0 \end{bmatrix}\right\} B , \qquad B^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} ,$$

and

$$\pi = (B^{-1})_4 = (1, -1, -1, -1) .$$

Determining the column to enter the basis

$$c_j = \pi(A^j - A^{k_i}) \quad \text{if} \quad A^j \varepsilon S_i , \quad A^j \text{ nonbasic.}$$

Beginning left to right,

$$c_1 = \pi(A^1 - A^3) = (1, -1, -1, -1) \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix} = -2 .$$

Therefore A^1 enters. $A^1 \varepsilon S_1$; and $\sigma = 1$.

Computing D^{1*} and $\underline{A^{1*}}$:

$$D^{1*} = B^{-1}(A^1 - A^3) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ -2 \end{bmatrix} ;$$

$$A^{1*} = \left[\begin{array}{c|c} I_{pxp} & -C_{pxm} \\ \hline 0_{mxp} & I_{mxm} \end{array}\right] \begin{bmatrix} 1 \\ \hline 1 \\ -1 \\ 2 \\ -2 \end{bmatrix} = I_{5x5} \begin{bmatrix} 1 \\ \hline 1 \\ -1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ \hline 1 \\ -1 \\ 2 \\ -2 \end{bmatrix} .$$

Note: $-C_{pxm} = (0,0,0,0)$ since columns 2,3,4,5 of $\underline{B}$ are from S_0 .

<u>Determining the column to leave the basis</u>

$$\Theta = \min_{A_i^{1*} > 0} \left\{ \frac{b_i}{A_i^{1*}} \right\} = \frac{b_r}{A_r^{1*}} = \min \left(\underset{\substack{\uparrow \\ A^3}}{\frac{1}{1}} , \underset{\substack{\uparrow \\ A^9}}{\frac{1}{1}} , \underset{\substack{\uparrow \\ A^{11}}}{\frac{2}{2}} \right) = 1 .$$

We choose to let A^9 leave the basis, $A^9 \in S_0$, $\rho = 0$.

<u>Updating B^{-1} and $\underline{b}$</u> A^9 is not a key column so we follow case 1. A^9 is presently the $1+1$st column of $\underline{B}$ so that $i_2 = 1$. Adjoin to B^{-1} the column D^{1*} and pivot on D_1^{1*} . Consequently

$$[D^{1*} \vdots B^{-1}] = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \\ -2 & 1 & -1 & -1 & -1 \end{bmatrix} \sim \underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 1 & -1 & -1 \end{bmatrix}}_{B^{-1} \text{ new}} .$$

The new values of the basic variables are:

$$b_1 \text{ new} = b_1 - \theta A_1^{1*} = 1 - (1\cdot 1) = 0 ;$$

$$b_2 \text{ new} = \theta = 1 = 1 ;$$

$$b_3 \text{ new} = b_3 - \theta A_3^{1*} = 1 - (1\cdot -1) = 2 ;$$

$$b_4 \text{ new} = b_4 - \theta A_4^{1*} = 2 - (1\cdot 2) = 0 ;$$

$$b_5 \text{ new} = b_5 - \theta A_5^{1*} = -4 - (1\cdot -2) = -2 .$$

Therefore the new basis and basic variables are

$$\underline{B} = (A^3 | A^1 A^{10} A^{11} A^0) \quad \text{and} \quad \underline{b} = (0,1,2,0,-2) .$$

Iteration 2

Determining the column to enter the basis   Computing the c_j's we have

$$\pi = (B^{-1})_4 = (1, 1, -1, -1) .$$

$$c_2 = \pi A^4 = (1, 1, -1, -1) \begin{pmatrix} 0 \\ -1 \\ 2 \\ 3 \end{pmatrix} = -6 .$$

Therefore A^2 enters the basis. $A^2 \in S_1$; $\sigma = 1$.

Computing D^{2*} and $\underline{A^{2*}}$:

$$D^{2*} = B^{-1}(A^2 - A^3) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5 \\ -6 \end{bmatrix} ;$$

$$\underline{A^{2*}} = \left[\begin{array}{c|cccc} 1 & -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} 1 \\ \hline -1 \\ 1 \\ 5 \\ -6 \end{array}\right] = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 5 \\ -6 \end{bmatrix} .$$

Notice that $-C_{pxm} = (-1,0,0,0)$; the -1 corresponds to $A^1 \in S_1$.

<u>Determining the column to leave the basis</u>

$$\theta = \min_{A_i^{2*} > 0} \left\{ \frac{b_i}{A_i^{2*}} \right\} = \frac{b_r}{A_r^{2*}} = \min \Big(\underset{\uparrow \atop A^3}{\frac{0}{2}} , \frac{2}{1} , \underset{\uparrow \atop A^{11}}{\frac{0}{5}} \Big) = 0 .$$

We choose to let A^3 leave the basis. $A^3 \in S_1$; $\rho = 1$.

<u>Updating B^{-1} and $\underline{b}$</u> A^3 is the key column of S_1 , and since A^2 and A^3 are both from the same set we follow Case 2b. Since A^1 is a nonkey member of S_1 that is a member of $\underline{B}$ we must change the key for S_1 from A^3 to A^1 as is described in Case 2a. A^1 is presently the $1+1$st column of $\underline{B}$ so that $i_2 = 1$. The key will be reassigned

by multiplying B by the following matrix T_1 :

$$T_1 = \begin{array}{cl} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{array}{l} \longleftarrow \text{ row } i_2 \\ \\ \\ \\ \end{array} \end{array} .$$

The -1 of row i_2 corresponds to $A^1 \in S_1$.

Since $T_1^{-1} = T_1$, then B^{-1} is replaced with $\tilde{B}^{-1} = T_1 B^{-1}$.

$$\tilde{B}^{-1} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} .$$

D^{2*} and A^{2*} must also be updated.

$\tilde{D}^{2*} = \tilde{B}^{-1}(A^2 - A^1)$, since A^1 is now the key for S_1 .

$$\tilde{D}^{2*} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 5 \\ -6 \end{bmatrix}$$

and

$$\underline{\tilde{A}^{2*}} = \left[\begin{array}{c|cccc} 1 & -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} 1 \\ \hline 2 \\ 1 \\ 5 \\ -6 \end{array}\right] = \left[\begin{array}{c} -1 \\ \hline 2 \\ 1 \\ 5 \\ -6 \end{array}\right] .$$

To update B^{-1}, adjoin $\tilde{D}^{2*}$ to $\tilde{B}^{-1}$ and pivot on $\tilde{D}_1^{2*}$.

$$[\tilde{D}^{2*} \vdots \tilde{B}^{-1}] = \left[\begin{array}{c|cccc} 2 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 5 & 0 & -2 & 0 & 1 \\ 6 & 1 & 1 & -1 & -1 \end{array}\right] \sim \underbrace{\left[\begin{array}{c|cccc} 1 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 3/2 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 \\ 0 & 1 & -2 & -1 & -1 \end{array}\right]}_{B^{-1} \text{ new}} .$$

The new values of the basic variables are:

b_1 new $= b_2 - \theta \tilde{A}_2^{2*} = 1 - (0 \cdot 2) = 1$, due to the key change;

b_2 new $= \theta = 0 = 0$;

b_3 new $= b_3 - \theta \tilde{A}_3^{2*} = 2 - (0 \cdot 1) = 2$;

b_4 new $= b_4 - \theta \tilde{A}_4^{2*} = 0 - (0 \cdot 5) = 0$;

b_5 new $= b_5 - \theta \tilde{A}_5^{2*} = -2 - (0 \cdot -6) = -2$.

Note: The above calculations are given in detail in case a similar situation occurred when $\theta \neq 0$. Therefore the new basis and basic variables are

$$\underline{B} = (A^1 | A^2 A^{10} A^{11} A^0) \quad \text{and} \quad \underline{b} = (1,0,2,0,-2) .$$

Iteration 3

Determining the column to enter the basis   Computing the c_j's we have

$$\pi = (B^{-1})_4 = (1,-2,-1,-1) .$$

$$c_4 = \pi A^4 = (1,-2,-1\ -1) \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} = 1 .$$

$A^4 \in S_0$, but since $1 > 0$, then A^4 does not enter the basis.

$$c_5 = \pi A^5 = (1,-2,-1,-1) \begin{pmatrix} 0 \\ -2 \\ -3 \\ 0 \end{pmatrix} = 7 .$$

$A^5 \in S_0$, however A^5 does not enter the basis since $7 > 0$.

$$c_6 = \pi A^6 = (1,-2,-1,-1) \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = -3 .$$

$A^6 \in S_0$, and A^6 enters the basis.

$A^6 \in S_0$; $\sigma = 0$.

Computing D^{6*} and $\underline{A^{6*}}$:

$$D^{6*} = B^{-1}(A^6) = \begin{bmatrix} 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 3/2 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 1 \\ 1 & -2 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 5/2 \\ \frac{1}{2} \\ -3 \end{bmatrix} ;$$

$$\underline{A^{6*}} = \left[\begin{array}{c|cccc} 1 & -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} 0 \\ \hline -\frac{1}{2} \\ 5/2 \\ \frac{1}{2} \\ -3 \end{array}\right] \text{, since } \sigma = 0 = \left[\begin{array}{c} \frac{1}{2} \\ \hline -\frac{1}{2} \\ 5/2 \\ \frac{1}{2} \\ -3 \end{array}\right]$$

Notice that $-C_{pxm} = (-1,0,0,0)$ the -1 corresponding to $A^2 \in S_1$.

<u>Determining the column to leave the basis</u>

$$\theta = \min_{A_i^{6*} > 0} \left\{ \frac{b_i}{A_i^{6*}} \right\} = \frac{b_r}{A_r^{6*}} = \min \left(\underset{\substack{\uparrow \\ A^1}}{\frac{1}{\frac{1}{2}}} , \underset{\substack{\uparrow \\ A^{10}}}{\frac{2}{5/2}} , \underset{\substack{\uparrow \\ A^{11}}}{\frac{0}{\frac{1}{2}}} \right) = 0 .$$

A^{11} leaves the basis. $A^{11} \in S_0$; $\rho = 0$.

<u>Updating B^{-1} and $\underline{b}$</u> Since A^{11} is not a key we follow Case 1. A^{11} is presently the $1+3$rd column of $\underline{B}$ so that $i_2 = 3$. Adjoin D^{6*} to B^{-1} and pivot on D_3^{6*}.

$$[D^{6*} \vdots B^{-1}] = \left[\begin{array}{c|cccc} -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 \\ 5/2 & 0 & 3/2 & 2 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 1 \\ -3 & 1 & -2 & -1 & -1 \end{array}\right] \sim \left[\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & -5 \\ 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 1 & -1 & 5 \end{array}\right] .$$

$$\underbrace{\qquad\qquad}_{B^{-1} \text{ new}}$$

The new values of the basic variables are:

$$b_1 \text{ new} = b_1 - \theta A_1^{6*} = 1 - (0 \cdot \tfrac{1}{2}) = 1 ;$$

$$b_2 \text{ new} = b_2 - \theta A_2^{6*} = 0 - (0 \cdot -\tfrac{1}{2}) = 0 ;$$

$$b_3 \text{ new} = b_3 - \theta A_3^{6*} = 2 - (0 \cdot 5/2) = 2 ;$$

$$b_4 \text{ new} = \theta = 0 = 0 ;$$

$$b_5 \text{ new} = b_5 - \theta A_5^{6*} = -2 - (0 \cdot -3) = -2 .$$

Therefore the new basis and values of the basic variables are

$$\underline{B} = [A^1 | A^2 A^{10} A^6 A^0] \quad \text{and} \quad \underline{b} = (0,0,2,0,-2) .$$

Iteration 4

Determining the column to enter the basis   Computing the c_j's we have

$$c_4 = \pi A^4 = (1,1,-1,5)\begin{pmatrix}0\\0\\0\\-1\end{pmatrix} = -5 .$$

$$A^4 \in S_0 .$$

Therefore A^4 enters the basis. $A^4 \in S_0$; $\sigma = 0$.

Compute D^{4*} and $\underline{A}^{4*}$:

$$D^{4*} = B^{-1}A_4 = \begin{bmatrix}0&0&0&1\\0&-1&1&-5\\0&1&0&2\\1&1&-1&5\end{bmatrix}\begin{bmatrix}0\\0\\0\\-1\end{bmatrix} = \begin{bmatrix}-1\\5\\-2\\-5\end{bmatrix} ;$$

$$\underline{A}^{4*} = \begin{bmatrix}0&-1&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}\begin{bmatrix}0\\-1\\5\\-2\\-5\end{bmatrix} \cdot \text{ since } \sigma = 0 = \begin{bmatrix}1\\-1\\5\\2\\-5\end{bmatrix}.$$

Determining the column to leave the basis

$$\theta = \min_{A_i^{4*} > 0} \left\{\frac{b_i}{A_i^{4*}}\right\} = \frac{b_r}{A_r^{4*}} = \min\left(\underset{\substack{\uparrow\\A^1}}{\frac{1}{1}} , \underset{\substack{\uparrow\\A^{10}}}{\frac{2}{5}}\right) = \frac{2}{5} .$$

Therefore A^{10} leaves the basis. $A^{10} \varepsilon S_0$; $\rho = 0$.

Updating the right-hand side This time we will update the right-hand side first, since we will have the optimal solution. The new values of the basic variables are:

$$b_1 \text{ new} = b_1 - \theta A_1^{4*} = 1 - (2/5 \cdot 1) = 3/5 ;$$

$$b_2 \text{ new} = b_2 - \theta A_2^{4*} = 0 - (2/5 \cdot -1) = 2/5 ;$$

$$b_3 \text{ new} = \theta = 2/5 = 2/5 ;$$

$$b_4 \text{ new} = b_4 - \theta A_4^{4*} = 0 - (2/5 \cdot -2) = 4/5 ;$$

$$b_5 \text{ new} = b_5 - \theta A_5^{4*} = -2 - (2/5 \cdot -5) = 0 .$$

Since $b_5 = 0$ we have the optimal solution. Therefore,

$$\underline{B} = [A^1 | A^2 A^4 A^6 A^0] \quad \text{and} \quad \underline{b} = [3/5, 2/5, 2/5, 4/5, 0] .$$

Note: If A^1 had been selected as the original key for S_1 , the lengthy iteration 2 would have been eliminated. We selected A^3 as the original key for S_1 in order to illustrate the calculations involved with a key change.

BIBLIOGRAPHY

Abrams, R. and Ben-Israel, A.
1968 A duality theorem for complex quadratic programming. Journal of Optimization Theory and Applications 4: 244-252.

Anderson, T.
1958 An introduction to multivariate statistical analysis. New York, New York, John Wiley and Sons, Inc.

Arrow, K. J. and Enthoven, A. C.
1961 Quasi-concave programming. Econometrica 29: 779-800.

Bartels, R. H.
1971 A stabilization of the simplex method. Numerische Mathematik 16: 414-434.

Bartels, R. H. and Golub, G. H.
1969 The simplex method of linear programming using L. U. decomposition. Communications ACM 12: 266-268.

Bartels, R. H., Golub, G. H. and Saunders, M. A.
1970 Numerical techniques in mathematical programming. Nonlinear Programming. New York, New York, Academic Press, Inc.

Bartels, R. H., Stoer, J. and Zenger, C.
1971 A realization of the simplex method based on triangular decomposition. Berlin-Heidelberg, Germany, Springer-Verlag.

Ben-Israel, A.
1969 Linear equations and inequalities on finite-dimensional, real, or complex, vector spaces: a unified theory. Journal of Mathematical Analysis and Applications 27: 367-389.

Ben-Israel, A., Charnes, A. and Kortanek, K.
1969 Duality and asymptotic solvability over cones. Bulletin of the American Mathematical Society 75: 318-324.

Berman, A. and Ben-Israel, A.
1971 Linear inequalities, mathematical programming, and matrix theory. Journal of Mathematical Programming 1: 291-300.

Björk, A. and Golub, G. H.
1967 Iterative refinement of linear least squares solutions by Householder transformations. BIT 7: 322-337.

Cooper, L. and Steinberg, D.
1970 Introduction to methods of optimization. Philadelphia, Pennsylvania, W. B. Saunders, Co.

Cottle, R. W.
1963a Symmetric dual quadratic programs. Quarterly of Applied Mathematics 21: 237-243.

Cottle, R. W.
1963b Theorem of Fritz John in mathematical programming. RAND Memorandum RM-3858-PR, Santa Monica, California.

Cottle, R. W. and Ferland, J. A.
1972 Matrix-theoretic criteria for the quasi-convexity and pseudo-convexity of quadratic functions. Journal of Linear Algebra and Applications 5: 123-136.

Cottle, R. W. and Sacher, R.
1973 On the solution of large, structured, linear complementary problems: I. Technical Report 73-74, Department of Operations Research, Stanford University.

Cottle, R. W., Habetler, G. J. and Lemke, C. E.
1967 Quadratic forms semi-definite over convex cones. Proceedings of the Sixth International Symposium in Mathematical Programming, Princeton University Press.

Dantzig, G. B.
1963 Linear programming and extensions. Princeton, New Jersey, Princeton University Press.

Dantzig, G. B. and Van Slyke, R. M.
1966 A generalized upper-bounded technique for linear programming. Proceedings of the IBM Scientific Computing Symposium on Combinatorial Problems, IBM, White Plains, New York.

Dantzig, G. B. and Van Slyke, R. M.
1967 Generalized upper bounding techniques. Journal of Computer and System Sciences 1: 213-226.

Dantzig, G. B., Eisenberg, E. and Cottle, R. W.
1965 Symmetric dual nonlinear programs. Pacific Journal of Mathematics 15: 809-812.

Dantzig, G. B., Wilson, R., Cottle, R. W. and Eaves, C.
1973 Systems Optimization Project: U. S. Atomic Energy Commission Report Code: SV-326PA18, Department of Operations Research, Stanford University.

Dorn, W. S.
1960 Duality in quadratic programming. Quarterly of Applied Mathematics 18: 155-160.

Dorn, W. S.
1961 Self-dual quadratic programs. Journal of the Society for Industrial and Applied Mathematics 9: 51-54.

Forrest, J. J. and Tomlin, J. A.
1972 Updated triangular factors of the basis to maintain sparsity in the product form simplex method. Journal of Mathematical Programming 2: 263-278.

Frank, M. and Wolfe, P.
1956 An algorithm for quadratic programming. Naval Research Logistics Quarterly 3: 95-110.

Geoffrion, A. M.
1970 Elements of large-scale mathematical programming. Management Science 16: 652-691.

Gould, F. and Tolle, J.
1972 Geometry of optimality conditions and constraint qualifications. Journal of Mathematical Programming 2: 1-18.

Guignard, M.
1969 Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. SIAM Journal on Control 7: 232-241.

Hanson, M. A. and Mond, B.
1967 Quadratic programming in complex space. Journal of Mathematical Analysis and Applications 23: 507-514.

Hanson, R. J. and Lawson, C. L.
1969 Extensions and applications of the Householder algorithm for solving linear least squares problems. Mathematics of Computation 23: 787-812.

Harris, P. M.
1973 Pivot selection methods of the DEVEX L. P. code. Journal of Mathematical Programming 5: 1-28.

Hemmerle, W.
1967 Statistical computations on a digital computer. Waltham, Massachusetts, Blaisdell Publishing Co.

Hildreth, C.
1954 Point estimates of ordinates of concave functions. JASA 49: 598-619.

Hildreth, C.
1957 A quadratic programming procedure. Naval Research Logistics Quarterly 4: 79-85.

Judge, G. and Takayama, T.
1966 Inequality restrictions in regression analysis. JASA 61: 161-181.

Kaul, R. N.
1965 An extension of generalized upper bounding techniques for linear programming. Report ORC 65-27, Operations Research Center, University of California at Berkeley.

Kempthorne, O. and Folks, L.
1971 Probability, statistics and data analysis. Ames, Iowa, The Iowa State University Press.

Krafft, O.
1970 Programming in statistics and probability. Nonlinear Programming. New York, New York, Academic Press, Inc.

Kretschmer, K. S.
1961 Programmes in paired spaces. Canadian Journal of Mathematics 13: 221-238.

Kuhn, H. W. and Tucker. A. W.
1951 Nonlinear programming. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, California.

Kunzi, H. and Krelle, W.
1966 Nonlinear programming. Waltham, Massachusetts, Blaisdell Publishing Co.

Laning, J. H. and Battin, R. H.
1956 Random processes in automatic control. New York, New York, McGraw-Hill Book Company, Inc.

Larsen, L. J.
1962 A modified inversion procedure for product form of the inverse linear programming codes. Communications ACM 5: 382-383.

Lasdon, L. S.
1970 Optimization theory for large systems. New York, New York, The Macmillan Co.

Lowell, M. and Prescott, E.
1970 Multiple regression with inequality constraints: pretesting bias, hypothesis testing and efficiency. JASA 65: 913-925.

Mangasarian, O. L.
1965 Pseudo-convex functions. SIAM Journal on Control 3: 281-290.

Mangasarian, O. L.
1969 Nonlinear programming. New York, New York, McGraw-Hill Book Company, Inc.

Mangasarian, O. L. and Fromovitz, S.
1967 The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. Journal of Mathematical Analysis and Applications 17: 37-47.

Mantel, N.
1969 Restricted least squares regression and convex quadratic programming. Technometrics 11: 763-773.

Mathematical Programming System/360
1975 Linear and separable programming-manual. IBM Report H20-0476-0.

Mond, B. and Cottle, R. W.
1966 Self-duality in mathematical programming. Journal of the Society for Industrial and Applied Mathematics 14: 420-423.

Mond, B. and Hanson, M. A.
1968 Symmetric duality for quadratic programming in complex space. Journal of Mathematical Analysis and Applications 23: 284-293.

Nelson, D. L., Lewis, T. O. and Boullion, T. C.
1971 A generalized inverse method for regression analysis with linear restrictions. Paper presented at the 1971 Joint Statistical Meetings, Fort Collins, Colorado.

O'Hagan, A.
1973 Bayes estimation of a convex quadratic. Biometrika 60: 565-571.

Orchard-Hays, W.
1968 Advanced linear programming computing techniques. New York, New York, McGraw-Hill Co.

Pyne, D.
1972 Duality in abstract mathematical programming with applications to statistical problems. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Raffin, C.
1968 Programmes linéaires d'appui d'un programme convexe, application avec conditions d'optimalité et à la dualité. Revue Francaise de Recherche Opérationnelle 13: 27-60.

Rosen, J. B.
1964 Primal partition programming for block diagonal matrices. Numerishe Mathematik 6: 250-260.

Slater, M.
1951 Lagrange multipliers revisited: a contribution to nonlinear programming. RAND Corporation Report RM-676, Santa Monica, California.

Sposito, V. A. and David, H. T.
1971 Saddlepoint optimality criteria of nonlinear programming problems over cones without differentiability. SIAM Journal on Applied Mathematics 20: 698-708.

Sposito, V. A. and David, H. T.
1972 A note on Farkas lemmas over cone domains. SIAM Journal on Applied Mathematics 22: 356-358.

Theil, H. and Van de Panne, C.
1960 Quadratic programming as an extension of classical quadratic programming. Management Science 7: 1-20.

Tomlin, J. A.
1972 Survey of computational methods for solving large scale systems. Technical Report 72-75, Operations Research House, Stanford University.

Varaiya, P. P.
1965 Nonlinear programming and optimal control. ERL Technical Memo M-129, University of California at Berkeley.

Wagner, H. M.
1959 Linear programming techniques for regression analysis. JASA 54: 206-212.

Wegner, P.
1963 Relations between multivariate statistics and mathematical programming. Applied Statistics 12: 146-150.

Wolfe, P.
1957 The simplex method for quadratic programming. RAND Report P-1205, Santa Monica, California.

Zellner, A.
1961 Linear regression with inequality constraints on the coefficients: An application of quadratic programming and linear decision rules. International Center for Management Science Report 6109 (MSN09), Rotterdam, England.

Zellner, A.
1963 Decision rules for economic forecasting. Econometrica 31: 111-130.

ACKNOWLEDGEMENTS

Dr. V. A. Sposito introduced me to the area of mathematical programming; without his assistance, patience and willingness to let me begin work on this dissertation while I was employed as a Des Moines school teacher, I might never have completed the Ph.D. program.

Various other faculty members deserve thanks. Without reservations, I wish to thank the support given to me by my committee members: Dr. D. F. Cox, Dr. H. T. David, Dr. W. A. Fuller, Dr. R. D. Hickman, and Dr. C. G. Maple. I thank Dr. Earl Heady for the funds contributed to the Numerical Analysis Section; Chapter V of this dissertation is a result of this support. I am indebted to Dr. Hickman and Dr. Fuller of the Survey Section and the "Des Moines River Data" for the support during my final year of study. Special thanks is due Department Chairmen Dr. T. A. Bancroft and Dr. H. A. David for supervising my NDEA Fellowship.

There are many other people who also deserve thanks. C. Asok, R. Carter and D. Dickey read parts and helped with the details of putting this dissertation together. C. Dunn typed much of the first draft and A. Jacobson typed the final copy. Finally, I must thank the many people of Snedecor Hall who supported my baking hobby, which was a necessary diversion while in graduate school.